OPTIMIZATION OF RESIDENTIAL MICROGRIDS USING BLACK-BOX MODELING METHODS

A Thesis
Presented to
The Academic Faculty

by

Christy Green

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Mechanical Engineering

Georgia Institute of Technology
August 2020

OPTIMIZATION OF RESIDENTIAL MICROGRIDS USING BLACK-BOX MODELING METHODS

Approved by:

Dr. Srinivas Garimella, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Berdinus Bras
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Richard Simmons
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: [June 22, 2020]

To my family

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to Dr. Srinivas Garimella, my advisor, for his encouragement and guidance throughout this process. He has challenged me to produce high-quality work and advocated on my behalf on numerous occasions. I am thankful for the opportunity he has provided me to be a part of the Sustainable Thermal Systems Laboratory. I would also like to thank Dr. Berdinus Bras and Dr. Richard Simmons for serving on my committee.

This work would also not have been possible without Steve Wilson, Tracy West, Pradeep Vitta, Justin Hill, and Jim Leverette of Southern Company Research & Development. I would like to thank them for funding this project and providing technical assistance along the way. I would also like to thank the current and former STSL members, particularly Bachir El Fil, Khoudor Keniar, Girish Kini, and Sriram Chandresekaran for their assistance and camaraderie during this process. I cannot imagine accomplishing this without them.

Last, but not in any way least, I would like to thank my parents, grandparents, sister, and friends for their support. They have been a constant source of encouragement and motivation and have kept me focused on my goal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Abbreviations

AI	Artificial intelligence
AMI	Advanced metering infrastructure
ANN	Artificial neural network
ARMA	Autoregressive moving average model
CNN	Convolutional neural network
COP	Coefficient of performance
CPP	Critical peak pricing
DSM	Demand-side management
DR	Demand response
HEMS	Home energy management system
HERS	Home energy rating system
HPWH	Heat pump water heater
HVAC	Heating, ventilation, air conditioning
IoT	Internet of things
MAE	Mean absolute error
NARX	Nonlinear autoregressive network with exogenous inputs
PLM	Peak load management
PV	Photovoltaic
RESNET	Residential energy services network
RMSE	Root mean square error
RTP	Real-time pricing

TCL	Thermostatically controlled load
TOU	Time of use
WH	Water heater
XGB	Extreme gradient boosting

Variables

c	Specific heat capacity
C	Cost of electricity
E	Energy
K	Cost function
m	Mass
N	Number
P	Power
Q	Thermal energy
t	Time
T	Temperature
U	Binary control variable
V	Volume
W	Work

Subscripts

air	Indoor air
br	Bedrooms
ER	Electric resistance
$grid$	Utility electric grid

unctrl	Uncontrolled loads
ref	Reference state
w	Water

Superscripts

max	Maximum
min	Minimum
SP	Set point

Greek symbols

η	Efficiency
ρ	Density
τ	Time interval
Φ	Penalty function
ω	Weight for objective function

SUMMARY

Electric utilities have traditionally sought to match demand with supply by curtailing the rate of power production at central power plants or purchasing electrical power from other utilities within the grid network. However, the supply side can be costly and inefficient to manage. Managing the demand side with demand response programs which incentivize utility customers to curtail energy consumption during peak hours may provide a more effective way of balancing the grid.

This thesis investigates the optimization of a residential microgrid to respond to demand response signals that increase the dispatchability of locally generated solar photovoltaic (PV) energy while reducing total energy cost. The study is conducted using data gathered from a 62-home neighborhood located in Birmingham, AL. Homes in the neighborhood are equipped with circuit-level energy monitoring, indoor and outdoor temperature and relative humidity sensors, and hot and cold water flow metering. Water heater and Heating, Ventilation and Air-conditioning (HVAC) operational data are provided in real time by interfacing with the proprietary control systems of the HVAC unit and water heater. One-minute interval energy data are also gathered from the local PV array and one-minute weather data are received from a weather station located within the neighborhood.

The HVAC system and water heater, which are the most significant residential electric loads, are designated as the controllable loads in this study. The HVAC system is a variable capacity, variable fan speed heat pump unit that serves multiple zones. The water heater is a hybrid heat pump model, which may use a heat pump, electric resistance

elements, or a synchronous combination of the two to heat water. A comparison of system identification method accuracy for these systems is made between black-box models and grey-box models found in the literature. The black-box system identification was performed using extreme gradient boosting (XGB). The use of black-box models allows for a generalized, data-driven modeling approach that does not require the user to supply specific information regarding system geometry or thermodynamic system parameters.

A multi-objective optimization problem is then developed with the objectives of minimizing energy cost and equipment degradation while maximizing the consumption of locally generated PV energy. Constraints are applied to ensure that water heater tank temperatures and indoor zone temperatures do not fall outside their respective designated temperature dead bands. The optimization is conducted using 15-minute intervals over an 8-hour time horizon, where an optimum value is found for each time interval with respect to all other intervals within the time horizon. The number of potential solutions is too great to perform an exhaustive search of solutions; therefore, the derivative-free optimization problem resulting from the use of black-box models is solved using a constrained genetic algorithm. Results are compared with the optimization of the same scenario posed as a linear programming problem formulated using grey-box models and optimized using a convex programming solver.

Modeling the HVAC system and water heater using extreme gradient boosting (XGB) resulted in model accuracies similar to or greater than those of the grey-box modeling methods. Optimization using the grey-box models in the objective function was performed using data from the months of August and January under two optimization scenarios. In the first scenario, no solar PV energy generation was included in the

optimization to isolate the response of appliances to time-of-use pricing signals. This scenario resulted in a maximum total energy savings of 5%, a maximum peak energy reduction of 29%, and a maximum total cost reduction of 8%. In the second scenario, solar PV generation was included in the optimization problem to determine the maximum amount of load shifting and energy savings possible with this problem formulation. The second scenario resulted in a maximum total energy savings of 12.9 %, a maximum peak energy reduction of 41.3%, and a maximum total cost reduction of 16.6%. Optimization results of the grey-box models were then compared with optimization results over the same time period using the black-box models, to benchmark the performance of the optimization using black-box models with the performance using grey-box models. The accuracy of the black-box models and the similarity of their optimization outcome to that of optimization using grey-box models demonstrates that this is a viable method for optimization that allows for an adaptable modeling method that may be useful in a large-scale demand response network.

CHAPTER 1. INTRODUCTION

Renewable energy consumption in the United States increased 87% from 2000 to 2018, with the largest contributors to this increase being wind and solar energy (EIA, 2019b). As renewable energy generation across the United States increases, the non-dispatchable nature of wind and solar energy is expected to present increasing challenges in grid balancing for grid operators (Laugs et al., 2020). Supply-side management alone may be insufficient in handling these issues, as was seen in the 2000-2001 California energy crisis, which may have been avoidable had the utility been capable of passing the high generation cost to consumers, resulting in self-curtailment of energy consumption (Faruqui, 2009). Passing the higher generation cost to the customer base in the form of price-responsive demand response is a form of demand-side management. Demand-side management has the potential to provide peak demand savings in excess of that generated by regulating the supply side alone (Mathieu, 2012). The major part of demand response (DR) program participants are typically residential customers, but only 44% of their potential peak energy savings is captured (EIA, 2019a). Automated, user-friendly DR solutions that allow direct load control may increase the amount of peak energy savings in the residential sector, as well as increase participation rates. The expansion of smart grid technologies allows for increased automation capabilities in home energy management systems (HEMS) for DR markets, which may encourage residential customer participation in DR programs.

In this thesis, an automated, data-driven method for optimizing the energy consumption of the water heater and HVAC system (the two largest electrical loads in a

residence) is developed for implementation in a residential DR program. This method aims to shift load out of peak demand periods by maximizing the consumption of locally generated solar PV energy and using thermal systems such as the HVAC system and water heater as virtual batteries (i.e., thermal energy storage systems) to facilitate peak load shifting.

1.1 Electricity Demand Curve

Hourly electricity demand in the United States follows a predictable daily pattern (ISO-NE, 2020). Figure 1.1 shows examples of the daily load curve for New England on days in July and December 2019 based on data from ISO-NE (2020). Representative plots for a cooling season in July (Figure 1.1a) and a heating season in December (Figure 1.1b) were selected to demonstrate seasonal variations. Features of the two curves include a peak period in the evening and a morning ramping-up period that is most pronounced in winter

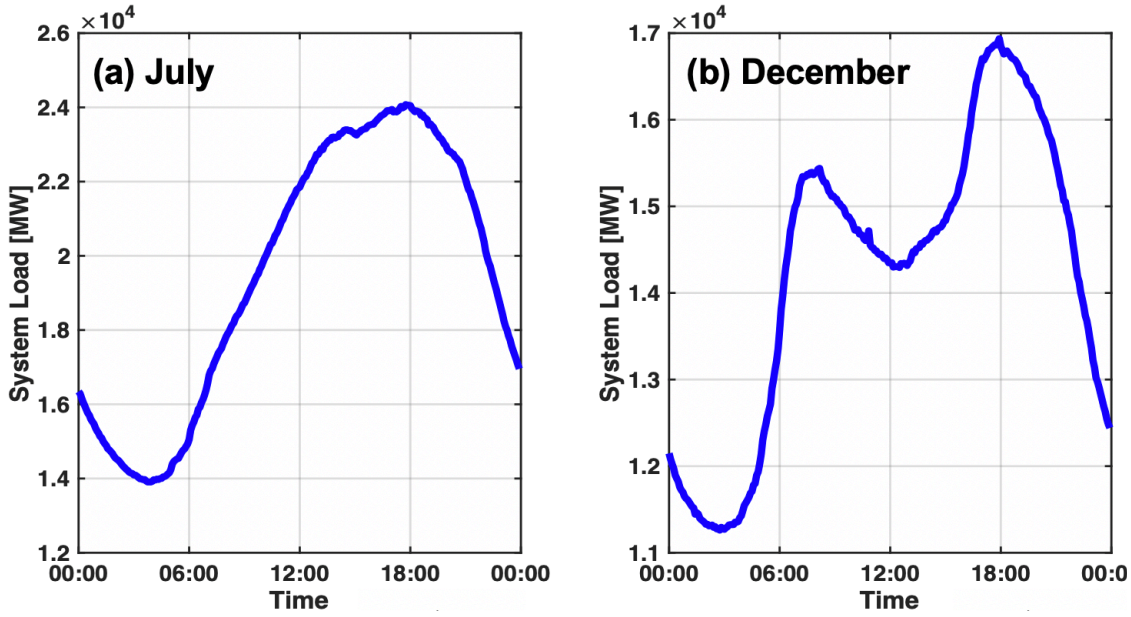

Figure 1.1: Electrical system load for ISO New England (ISO-NE, 2020) in (a) July, (b) December

months, especially in residences that use electric heating. The morning ramping period typically occurs between 5:00 AM and 8:00 AM, and puts significant stress on power systems. The peak period occurs in the evening hours, typically between 2:00 PM and 8:00 PM (EIA, 2019a). System operators must match generation and supply in real time, which requires dispatching or curtailing generation quickly.

1.2 Impact of Non-Dispatchable Generation on the Grid

The term non-dispatchable describes an energy generation system that cannot be expected to provide power on demand. Solar photovoltaic (PV) and wind energy systems are two examples of such systems. Both solar PV and wind energy follow seasonal and hourly patterns of fluctuation in generation capacity, with solar PV peaking during mid-day and wind energy systems typically peaking during night hours. Denholm et al. (2008) explored the energy production cost for high levels of solar PV generation. They provide a graphical representation of the magnitude of load shape impacts for a range of PV penetration scenarios from data provided by the Western Electricity Coordinating Council (WECC) in Figure 1.2. For higher levels of PV penetration, a much steeper ramp is observed between mid-day and the evening peak. This shape is commonly known as the "duck curve". The curve shows the energy demand on the electrical grid. The trough of the duck curve coincides with the peak daily period of PV energy production, thus the increase in ramping slope for higher levels of PV penetration. The U.S. Energy Information Administration predicts that renewable energy will provide 31% of energy in the U.S. by 2050, with 48% of that being generated by solar PV and 25% by wind (EIA, 2019b). With 30% PV penetration, the duck curve would be expected to be even more pronounced than the shapes shown in Figure 1.2.

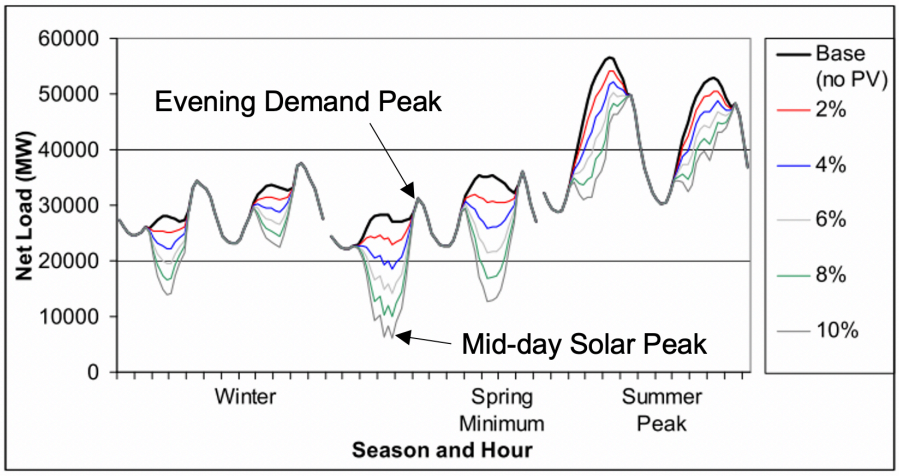

Figure 1.2: Load shapes in California for various WECC PV penetration scenarios (Denholm et al., 2008)

1.3 Demand Response

The challenges of generating and storing electricity, coupled with the high variability of electricity demand, result in problems of demand and supply matching that may be difficult to solve with supply-side management alone. Demand response programs were developed to augment supply-side management by managing the demand side of the electric grid. Demand response is defined by the U.S. Department of Energy as the deviation in electricity consumption by end-use consumers from their normal consumption patterns in response to time-based rate changes or incentivized programs to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized (DOE, 2006). Demand response allows electric utilities to manage electric loads in an effort to reduce or "smooth" peak periods in electrical demand. Table 1.1 provides the annual potential for peak demand savings in each of the North American Electric Reliability Corporation (NERC) regions. Potential peak savings are the total demand savings that may be achieved during a peak period if all demand response is called.

Table 1.1: Potential Peak Savings From Retail Demand Response Programs by NERC Region (FERC, 2018)

Region	Annual Potential Peak Demand Savings (MW)		Year-on-Year Change	
	2015	2016	MW	%
AK	27.0	27.0	0.0	0.0
FRCC	3,246.5	3,259.4	12.9	0.4
HI	35.2	33.5	-1.7	-4.8
MRO	4,508.9	5,231.3	722.4	16.0
NPCC	787.4	1,120.2	332.4	42.3
RF	5,372.2	5,505.1	132.9	2.5
SERC	9,259.1	8,265.6	-993.5	-10.7
SPP RE	1,922.7	5,004.4	3,081.7	160.3
Texas RE	696.4	773.3	76.9	11.0
WECC	7,019.2	6,625.3	-393.9	-5.6
Unspecified	0.0	79.0	79.0	--
Total	**32,874.6**	**35,924.1**	**3,049.5**	**9.3**

The SERC region, which covers much of the southeastern United States, has the highest potential for peak demand savings despite the 10.7% reduction in potential from 2015 to 2016 (FERC, 2018).

There are two primary types of demand response: price-responsive programs, and incentive-based programs. There are several methods for achieving price-responsive demand, including real-time pricing (RTP), time-of-use (TOU) rates, interruptible rates, and critical peak pricing (CPP) (Borenstein et al., 2002). RTP rate plans charge customers different prices hourly, with the hourly pricing varying from day to day. Customers may hedge energy sold under an RTP rate plan by negotiating prices for a given amount of energy ahead of time using a long-term contract to reduce the amount of energy purchased during peak RTP rate periods. TOU rates, which are the most common form of price-responsive demand response, vary the price of electricity for set blocks of time. For example, the hours of 2:00 PM to 7:00 PM may be considered peak-use hours during which

customers would be required to pay a higher price per kilowatt-hour than during off-peak hours. This rate plan also typically carries a peak demand charge that is based on the peak kilowatt demand of the customer during a given period, and is based only on the customer's energy consumption during that period, not the fluctuation of generation cost or wholesale prices. Interruptible pricing is nearly constant, with drastically higher prices charged during shortage periods in which customers subscribing to this plan are asked to cease electricity use. CPP plans are a blend of the previously mentioned approaches, consisting of a TOU rate structure combined with interruptible pricing periods. The utility is usually limited in the number of CPP events it may call in a year.

Incentive-based programs include direct load control (DLC), demand bidding/buyback programs, emergency demand response programs, capacity market programs, and ancillary services market programs (DOE, 2006). In DLC programs, the utility has control over a customer's electrical equipment, such as an air conditioner or water heater, and may cycle a system on or off with short notice to regulate demand. In demand bidding/buyback programs, customers provide bids to curtail energy consumption based on wholesale market prices. Demand bidding/buyback is typically offered to utility customers whose demand is >1 MW. In emergency DR programs, payments are given to customers that curtail energy consumption during grid emergencies. In a capacity market program, load curtailment offered by customers is treated as system capacity that replaces generation resources. This work considers a scenario in which direct load control is used in combination with time-of-use rates and PV, and weather forecasts to optimally schedule appliance operation.

1.4 Residential Energy Consumption

Residential energy consumption typically accounts for approximately 40% of the total energy consumption in the U.S., but during peak hours, it is closer to 50% (Asadinejad et al., 2018). For this reason, the residential energy market is a particularly attractive target for demand response programs. The challenge to DR in residential markets lies in the reliability of the demand response resources being called. For a price-responsive demand control method to be used, a metering infrastructure that allows the utility to determine the time at which energy is consumed by the premises must be in place. Advanced metering infrastructure (AMI) meters have replaced manual meters in many homes in the United States. At present, there are approximately 86.8 million AMI installations in the United States, and 88% of these are residential customers (EIA, 2019a). These meters typically allow the utility to remotely measure the energy consumption of a residence in 15- to 60-minute increments. This is a one-way communication from the meter to the utility; however, some meters have the ability to provide two-way communication in which the utility may send price signals to the meter as well as receive energy consumption data. AMI infrastructure alone allows the utility to offer TOU rates that align with peak demand periods, but coupling this capability with a HEMS that is capable of executing automated direct load control events would further enhance DR peak energy savings.

Internet of Things (IoT) devices, i.e., Internet-enabled devices, allow users to remotely control and schedule many devices in the home. As the range of IoT devices has expanded to include most major appliances, in addition to home security and entertainment devices, HEMS have become more common. While many current HEMS only provide

energy monitoring capabilities, there is an opportunity to use such technology for load forecasting and automated direct load control.

1.5 Scope of Present Work

The present work seeks to develop an optimization strategy for residential energy consumption that would help facilitate the introduction of automated demand response. Data-driven black-box models are developed for each thermostatically controllable load to provide a generalized model that may be applied to a variety of similar systems. These models are compared with grey-box models, which use a combination of physics-based and statistical modeling methods that were developed from models found in the literature. An optimization algorithm is then developed using these models with a genetic algorithm developed for time-series optimization. The objectives of the optimization are as follows:

- Reduce electricity cost for the residential customer
- Maximize the local dispatch of on-site solar PV energy
- Maintain customer comfort
- Shift electrical load out of peak demand periods

1.6 Thesis Organization

This thesis is organized into the following chapters:

Chapter Two presents a comprehensive review of the literature on home energy management systems, thermal system modeling, and optimization methods.

Chapter Three describes the neighborhood test site and data acquisition equipment.

Chapter Four describes the framework for the optimization of the hybrid heat pump water heater and HVAC system.

Chapter Five describes modeling of the thermal systems within the residences. Grey-box and black-box modeling methods are presented.

Chapter Six presents the optimization results for both grey-box and black-box modeling methods.

Chapter Seven presents the conclusions and recommendations for future work.

CHAPTER 2. LITERATURE REVIEW

Home energy management systems (HEMS) are essential to reduce energy consumption and redistribute energy as part of a utility-facilitated demand-side management (DSM) program (Shareef et al., 2018). A survey of HEMS-related literature, including the HEMS infrastructure, demand response program types, scheduling techniques, and modeling methods for home appliances is presented in this chapter. HEMS are vital to the optimization of the energy consumption and participation of a home in demand response activities, as they provide a means to measure and control connected devices within the home. Figure 2.1 presents a schematic of a HEMS, which shows the relation between each device or data acquisition unit. Shareef et al. (2018) provide a survey of HEMS and conclude that the future of HEMS lies in the further development of artificial

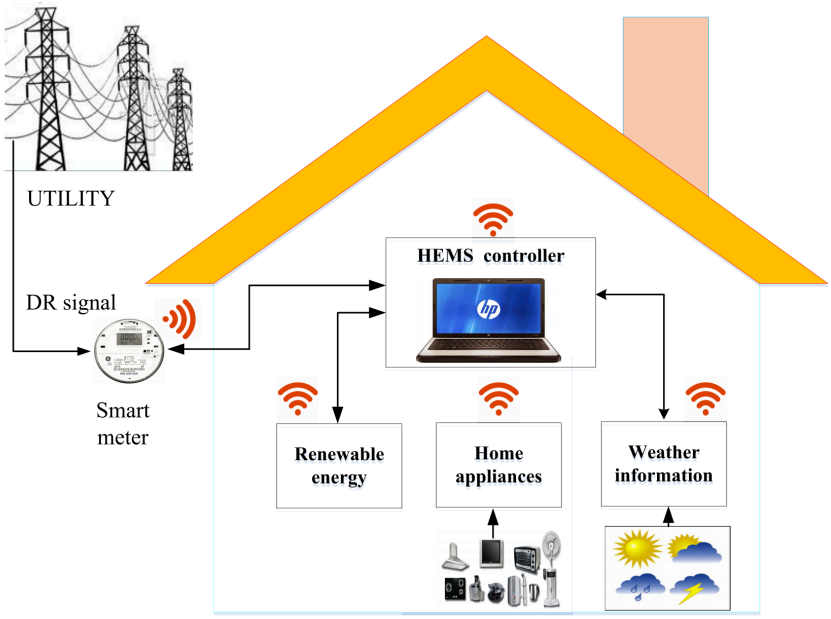

Figure 2.1: Flow diagram of HEMS interactions (Shareef et al., 2018)

intelligence (AI) systems to remove the need for end user involvement in the energy management system. Ideally, AI systems would be capable of developing data-driven models used for energy optimization, with minimal inputs from the end user. Leitão et al. (2020) conduct a review of the available HEMS technologies and conclude that coordinating distributed renewable energy generation with centralized power generation will be a critical contribution of HEMS in the future.

2.1 Energy Management in Buildings

2.1.1 Demand-Side Management

Palensky and Dietrich (2011) categorize DSM programs based on end user impact and timing as energy efficiency, time of use (TOU), demand response (DR), or spinning reserve. Energy efficiency programs encourage consumers to take measures that improve the efficiency of existing equipment, such as exchanging an inefficient space conditioning system for a more efficient one or adding more insulation to the building shell. TOU programs penalize energy consumption during certain periods by increasing the cost of energy during the period. Spinning reserve techniques adjust loads on the customer side in response to changes in grid frequency. DR programs influence energy consumption patterns by changing the price of electricity or providing incentive payments. Table 2.1 provides a comparison of residential DSM programs, highlighting the advantages and disadvantages of each. The necessity for action on the part of the consumer for each of the DSM programs is a disadvantage shared by all programs listed. This disadvantage may be overcome by using a HEMS with automated decision making.

Table 2.1: Residential DSM program comparison, adapted from Shareef et al. (2018)

Residential DSM program	Response and activation period	Advantage	Disadvantage
Real-time pricing	Electricity prices vary hourly for the end user, with pricing schedules updated daily	End user can reduce the electricity cost based on the price change	Customers who want to reduce electricity bill should instantaneously respond
Time of use pricing	Electricity prices vary hourly for the end user, according to a peak load schedule	Tariff prices are high during on-peak and low during off-peak, thereby encouraging customers to shift loads to reduce costs	Tariff displays one price change with respect to time for all customers and following the prices is compulsory for end users.
Critical peak pricing	Electricity prices for the end user may change at any time	End users receive a notification for a short period to earn a discount	End users should curtail or shift home devices for a certain period
Direct load program	Electricity prices for the utility may change at any time	The utility offers special discounts for shifting appliance loads	Utility curtails or shifts appliance power consumption to balance grid power with authorization from end users
Curtailable programs	Electricity prices for the end user may change at any time	End users respond within a limited period of time to obtain discount rates	End users curtail or shift appliance energy consumption for certain periods of time
Demand bidding	Electricity prices for the end user may change at any time	The utility offers special discounts for shifting appliance loads	End users curtail or shift appliance energy consumption for certain periods of time

2.1.2 Energy Billing Schemes

With advances in smart metering, electric utilities may offer a variety of rate plans that were not formerly available with manually-read electric meters to residential

customers. Bayram and Ustun (2017) and Rasheed et al. (2015) describe five main billing approaches: all-in-rate, TOU, critical peak pricing (CPP), real time pricing (RTP), and inclining block rate (IBR). The all-in-rate bills customers at a static rate throughout the day. TOU pricing divides the day into several periods and charges a fixed rate during each period. The periods and rates charged in a TOU scheme typically change seasonally. CPP billing schemes define a peak rate that customers are charged during critical grid times in which the grid is overloaded. RTP prices are continuously updated to reflect generation cost, typically on an hourly or daily basis.

2.2 Household Appliance Models and Scheduling Techniques

Accurate modeling of appliances within the household is necessary to provide meaningful estimations of the energy consumption for various scheduling schemes. Beaudin and Zareipour (2015) classify household appliances based on their load management possibilities in HEMS. The six categories identified are described in Table 2.2. All appliances may be classified as controllable or uncontrollable; the controllable appliances may be further classified as curtailable loads, uninterruptible loads, interruptible loads, regulating loads, or energy storage. The assignment of appliances to categories (discussed in Table 2.2) may vary according to the time scale on which management takes place. For example, a refrigerator may be considered interruptible when controlled at 5-minute intervals but uninterruptible on a 60-minute timescale. Some appliances may be dually classified; for example, a water heater is a regulated load, but may be used to store thermal energy.

Table 2.2: Appliance categories, adapted from Beaudin and Zareipour (2015)

Appliance Category	Description
Uncontrollable loads	Loads that cannot be altered or rescheduled by HEMS
Curtailable loads	Loads that can be adjusted mid-operation with little or no impact to residents' comfort
Uninterruptible loads	A start time may be specified for this appliance, but a complete cycle must be run once started
Interruptible loads	Can be interrupted and resumed at any time
Regulating loads	Appliance operational states are dictated by maintaining a given reference condition defined by residents or the HEMS (e.g., HVAC systems)
Energy storage	Appliances that store energy, such as external batteries

HEMS may provide a reduction in end use energy consumption or a reduction in peak energy demand by scheduling appliances in a way that meets occupants' comfort requirements while consuming the minimal amount of energy to do so. A common control strategy involves minimizing energy cost using a dynamic tariff while also minimizing peak demand (Marzband et al., 2017). A variety of control scheduling, modeling, and optimization methods have been investigated. Some commonly used methods include time series building energy modeling, artificial intelligence techniques, mathematical optimization, and heuristics and metaheuristics (Shareef et al., 2018), (Leitão et al., 2020).

For time series building energy modeling, predictive models are developed for all appliances in the house, and then the optimization is conducted over a specified time period, i.e., time horizon (Ahmad et al., 2018). As a result, the operational period of an appliance is optimized not only with respect to other synchronous appliance loads, but also to loads in all time steps across the time horizon. Statistical or machine learning approaches may be considered for use in these types of scheduling problems. Conventional statistical methods such as linear regression and autoregressive moving average model (ARMA) may

be used when the predicted load is linear (Shareef et al., 2018). In Jin et al. (2017), a linear model of a building energy system is presented and optimized using a convex linear solver. Simplified linear models and system identification methods are used to develop the models. When the load is nonlinear, artificial intelligence methods such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and restricted Boltzmann machines may be used. Amarasinghe et al. (2017) obtained similarly accurate results when they compared CNNs to support vector machine (SVM), LSTM, and Factor Restricted Boltzmann Machines (FCRBM). Additionally, Amarasinghe et al. (2017) compared various types of LSTM architectures for building energy modeling and determined that the LSTM performed better on 60-minute data than 5-minute data. They also found that there was little variation in performance across the various LSTM architectures for this task. Cui et al. (2019) developed various data-driven models for the prediction of indoor temperature response to a given HVAC electrical load and found that a relatively new type of predictive model, extreme gradient boosting (XGB), was the best for predicting indoor temperatures.

Appliance scheduling may be performed using a variety of methods, including mathematical optimization and heuristic/metaheuristic methods. Mathematical optimization finds feasible solutions that maximize or minimize a given objective function that may be subject to constraints (Jin et al., 2017). It is a deterministic optimization approach, which converges to a global minimum or maximum. Heuristic or metaheuristic methods do not guarantee a global minimum or maximum, but are useful for large problems in which mathematical optimization may not be reasonable due to intense computational requirements (Alajmi and Wright, 2014).

2.2.1 Optimization Objectives

The scheduling problem is usually formulated with a combination of multiple objectives (Jin et al., 2017). Some commonly used objectives include reduction of the customer's electricity bill, reduced distribution system losses, minimization of peak load, minimization of carbon emissions, maximizing customer comfort, and increasing social welfare. Table 2.3 provides a list of several studies and their respective optimization objectives. The electricity bill objective aims to minimize the cost of electricity to the user. Distribution system losses seeks to minimize distribution losses in the electrical grid by optimally scheduling generation sources along the power line. The peak load objective tries to minimize the amount of energy consumed during peak periods by curtailing appliance operation or interrupting appliance operation. The goal of the carbon emissions optimization is to minimize the amount of carbon produced by the optimized generation and consumption mix. The customer comfort objective seeks to maintain customer comfort by reducing the amount of load interruptions and ensuring that thermostatically controlled loads (TCLs) are maintained within the customer's comfort range. The social welfare objective seeks to improve the welfare of an entire community of participants, rather than a single end user, for example, by using game theory in a transactive energy environment to reduce energy cost collectively for all participants.

2.3 Research Opportunities

Opportunities exist for further research on the improvement of AI used in HEMS and the integration of renewable energy generation in the HEMS framework. AI may be used to reduce or eliminate customer inputs to the HEMS, resulting in greater energy

captured for a given objective. Integration of renewable energy on an individual or community scale allows for the maximum dispatchability of renewable energy resources and increased grid reliability. Additionally, integration of renewable energy assists in load shifting for peak load management by quantifying the amount of renewable energy forecasted to be available, and balancing it with the demand. This work investigates the potential for data-driven AI models of controllable household appliances that would minimize user inputs. Furthermore, the benefits and challenges associated with adding renewable energy generation into the scheduling problem are investigated. The project is divided into the following objectives/tasks:

- Linear mathematical modeling and optimization for HEMS to reduce peak demand, minimize energy consumption, and increase the dispatchability of locally generated solar PV energy while maintaining customer comfort,
- Data-driven modeling and metaheuristic optimization for HEMS, which satisfy the previously mentioned objectives.

Table 2.3: Optimization objectives

Paper	Electricity Bill	Distribution System Losses	Peak Load	Carbon Emissions	Customer Comfort	Social Welfare
Chavali et al. (2014)	✓					✓
Lokeshgupta and Sivasubramani (2019)	✓		✓			
Golmohamadi et al. (2019a)	✓					
Anvari-Moghaddam et al. (2015)	✓				✓	
Golmohamadi et al. (2019b)	✓					
Li et al. (2011)						✓
Miceli (2013)		✓				
Liu et al. (2014)			✓			
Bayram and Ustun (2017)						✓
Samadi et al. (2013)			✓			
Leitão et al. (2018)	✓					
Liu et al. (2014)			✓			
Nguyen and Le (2014)	✓					
Wu et al. (2016)	✓					
Sortomme et al. (2011)		✓				
Deilami et al. (2011)		✓				
Jin et al. (2017)	✓			✓	✓	
Saber and Venayagamoorthy (2011)				✓		
Bamdad et al. (2017)			✓			
Liao et al. (2018)			✓			
Verhelst et al. (2012)					✓	
Marzband et al. (2017)	✓					
Lampropoulos et al. (2019)	✓					
Adika and Wang (2014)			✓			
Bakker et al. (2010)				✓		
Chen et al. (2014)						✓
Ahmad et al. (2019)			✓	✓		

CHAPTER 3. TEST SITE

The test site, pictured in Figure 3.1, consists of a neighborhood of sixty-two single-family residences located in Birmingham, Alabama. The houses receive a portion of their energy from a community-level microgrid that consists of a photovoltaic (PV) array, natural gas generator, and a battery energy storage system. Any electrical power needed in excess of what is generated by the microgrid is provided from the local utility electrical grid. The houses are equipped with smart appliances that aid in the implementation of peak load management (PLM) and demand response (DR) events.

3.1 House Characteristics

The houses were constructed to maximize energy efficiency as a primary goal. Energy efficient construction measures taken include the use of high-resistance insulation,

Figure 3.1: Smart neighborhood and distributed generation (Southern Company Services ©)

double pane windows, and energy recovery ventilation. The houses in this study have Home Energy Rating System (HERS) ratings between 40 and 50, which is a 50-60% improvement over the Residential Energy Services Network (RESNET) reference score of 100 (RESNET, 2019). The HERS score is a metric that may be used to compare expected energy consumption between houses. A house with an HERS index score of 100 is the standard reference for a newly constructed house; a score less than 100 indicates energy savings. The percent savings indicated by the HERS index may be calculated using Equation (3.1).

$$\% \text{ Energy Reduction} = 100 - \text{HERS Rating} \tag{3.1}$$

The exact method of calculating the HERS score is proprietary information of RESNET. It is calculated using information typically used in developing building energy models, including insulation R-values, window U-values, house design characteristics, and thermal systems such as the HVAC system and water heater. The house audit is then compared with the HERS index new-construction reference home to designate an HERS rating.

The houses range in size from 1885 ft^2 (175.12 m^2) to 2876 ft^2 (267.19 m^2). Each house has one to three HVAC zones (depicted in Figure 3.2). The typical house size

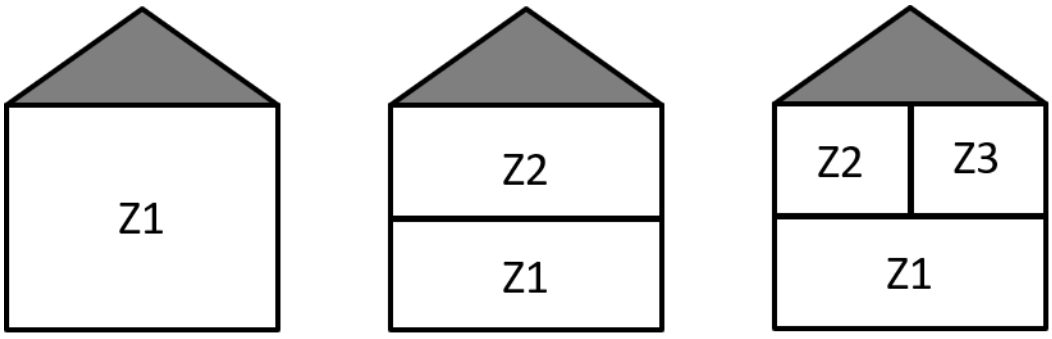

Figure 3.2: Schematic of HVAC zones

according to Building America House Simulation Protocol (BAHSP) 2014 is 2400 ft^2 with three bedrooms, which is approximately the median value for the houses in this neighborhood. According to BAHSP, the number of occupants can be estimated using Equation (3.2).

$$N_{occ} = 0.59N_{br} + 0.87 \tag{3.2}$$

where N_{occ} is the number of occupants and N_{br} is the number of bedrooms. The characteristics of each house type are given in Table 3.1, while the house construction characteristics are given in Table 3.2.

Table 3.1: House type characteristics, specified in construction documents

House Type	Quantity	Area (ft^2)	Area (m^2)	Number Bedrooms	Number Bathrooms	HVAC Zones
A	1	1885	175.1	3	2	1
B	13	2023	187.9	3	2.5	2
C	12	2467	229.2	4	2.5	2
D	29	2876	267.2	4	3	2-3
E	7	2486	231.0	4	3	2

Table 3.2: Construction characteristics, specified in construction documents

Component	R-Value ($h\ ft^2\ ^oF\ BTU^{-1}$)	R-Value ($m^2\ K\ W^{-1}$)	U-Value ($BTU\ h^{-1}\ ft^2\ ^oF^{-1}$)	U-Value ($W\ m^{-2}\ K^{-1}$)
Window	-	-	0.35	0.50
Insulation (Exterior Wall)	13	2.29	-	-
Insulation (Roof)	30	5.28	-	-

3.2 Local Climate

Birmingham, AL is classified as a Mixed-Humid climate as defined by the Department of Energy Building America climate region guide (Hendron and Engebrecht,

2010). The mixed-humid region is a region that receives over 20 in. (50 cm) of precipitation annually, has ≤ 5400 annual heating degree days (65ºF (18.3ºC) basis), and an average monthly outdoor temperature in winter that is ≤ 45ºF (7ºC) (EERE, 2019). For the two years monitored in this study (2018-2019), all criteria of the mixed-humid climate were met. The seasons are defined using the local electric utility's time-of-use (TOU) seasonal designation, given in Table 3.3.

Table 3.3: Season designation

Season	Months	Peak Hours	Off-Peak [¢ kWh^{-1}]	On-Peak [¢ kWh^{-1}]
Winter	November - March	05:00 – 09:00	6.954	8.954
Spring	April - May	-	-	-
Summer	June - September	13:00 – 19:00	6.954	26.954
Autumn	October	-	-	-

The most severe energy peaks occur at the hottest and coldest times of day due to the significance of the space-conditioning load. These periods occur in summer and winter months and are designated as time-of-use or "time advantage" hours for electric utility billing purposes, during which the cost of electricity is much higher than that of non-TOU hours. The hours of these peak periods and the corresponding rates are given in Table 3.3. The rates during on-peak and off-peak periods vary seasonally, with summer having the highest on-peak cost.

3.3 Appliances

Highly efficient smart appliances were selected for the houses in the neighborhood. Manufacturer and model number information for these and other appliances in the houses are given in Table 3.4. Further description of the controllable appliances is provided in Chapter 5. The water heater is an 80-gallon (0.30 m^3) hybrid heat pump unit manufactured

Table 3.4: Appliances installed in the neighborhood

Appliance	Manufacturer	Model Number
Water heater	Rheem	PROPH65 T2 RH350 DCB
HVAC unit	Carrier	24VNA0 (2-5 tons)
Refrigerator	Samsung	RF256BEAESG
Clothes washer	Samsung	WED9620HC
Clothes dryer	Samsung	WFW9620HBK
Dishwasher	Whirlpool	WDTA75SAHZ

by Rheem®. The HVAC systems are continuously variable heat pump units manufactured by Carrier®.

3.4 Site Monitoring

The houses are equipped with circuit-level energy monitoring that captures the energy consumption of appliances on each circuit as well as the total energy consumption of the house. Temperature and relative humidity sensors are located in each zone, in addition to the temperature readings taken at each zone's thermostat. Temperature and relative humidity are also measured in the unconditioned garage and attic spaces. The uncertainties of the sensors used in this study are provided in Table 3.5. An on-site weather station measures irradiance, ambient temperature, relative humidity, and wind speed.

Table 3.5: Sensor uncertainty

Parameter	Units	Sensor	Accuracy
Power	W	SiteSage Residential Power Monitor	±2%
Indoor temperature	ºC	PowerWise inSense	±0.5
Indoor relative humidity	%	SiteSage RH	±3%
Cold water volumetric flow rate	gal min^{-1}	SeaMetrics MJE	±0.25
Hot water volumetric flow rate	gal min^{-1}	SeaMetrics MJE	±0.25
Water heater reservoir temperature (upper and lower)	ºC	Thermistor	±0.5

3.5 House Energy Consumption

The seasonal mean energy consumption for the water heater and HVAC system in this neighborhood are plotted in Figure 3.3 and Figure 3.4, respectively. The trend of the power curve for the heat pump water heater remains similar seasonally as compared to that of the HVAC system that shifts its peak significantly from the heating to cooling season.

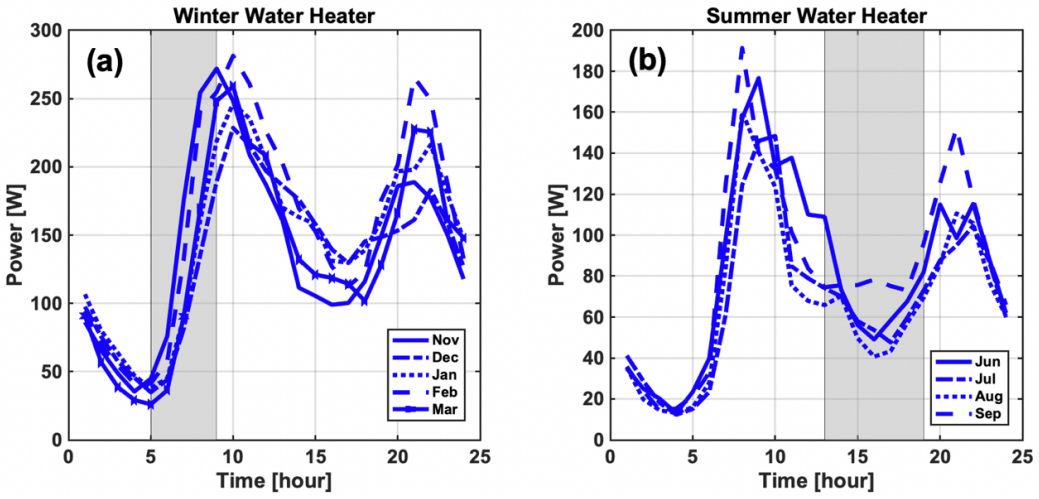

Figure 3.3: Mean hourly water heater energy consumption, (a) winter, (b) summer

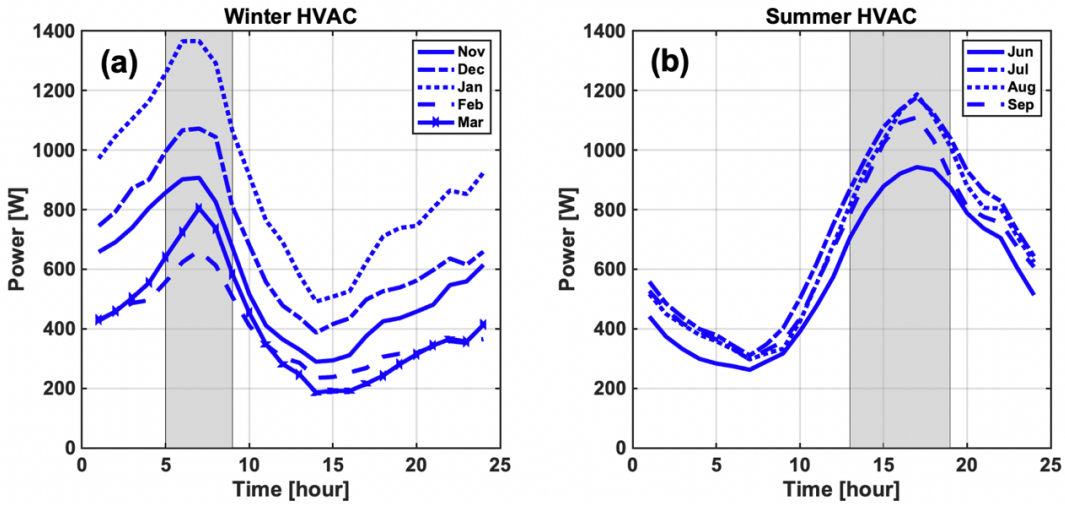

Figure 3.4: Mean hourly HVAC energy consumption, (a) winter, (b) summer

The total potential load shifting per house for the winter and summer periods was estimated using the mean energy consumption across the neighborhood for each appliance during each hour of the peak use period and is given in Table 3.6. The mean of energy consumption was calculated using Equation (3.3):

$$E_s = \sum_{h=1}^{H} \bar{P}_h t \tag{3.3}$$

where E_s is total amount of energy available for load shifting, t is the time interval, and $\bar{P}_h$ is given by Equation (3.4):

$$\bar{P}_h = \frac{\sum_{i=1}^{60} P_i}{60 \text{ min}} \tag{3.4}$$

The HVAC system has a much greater amount of energy available for shifting, but it subject to more stringent temperature constraints due to its effect on user comfort.

3.6 Concurrent Research Efforts and Monitoring Challenges

The neighborhood is an active test bed for a variety of research organizations. During the data collection period, testing was in progress for a control scheme that involved load control on a week-on, week-off schedule. Thus, data used in this study were filtered to remove the periods in which active testing was taking place within the neighborhood. This data filtering was performed using a test schedule of the test dates that was verified

Table 3.6: Energy available for load shifting (peak energy)

Season	Total House Energy [kWh]	Controllable Appliance Energy [kWh]	
		Water Heater	HVAC
Winter	6.37	0.59	3.53
Summer	9.78	0.39	6.10

by checking for constant set points and modes on each thermostatically controlled load during the inactive test weeks. In addition to interference from other testing activities, network connectivity, and sensor issues led to frequent periods of partially missing data. One house in the neighborhood was completely unaffected by external testing and had a complete data set for the desired time period. For this reason, this work focuses on methods for extending the modeling and optimization of controllable appliances to a wider group of homes, rather than focusing on the modeling and optimization of all houses in this neighborhood. Data-driven methods for thermal systems modeling are investigated to develop a set of models that may be applied to other houses within the neighborhood or houses in other locations.

3.7 Site Comparison

It is important to note differences between this site and others when drawing conclusions from the optimization analysis. Variables such as geographic location, house construction, microgrid specifications, and local energy policies may influence results.

First, we may consider the variability of solar PV generation from region to region. System Advisor Model (SAM), a tool developed by NREL to model renewable energy systems, was used to model the effect of geographic location on PV energy production. A city corresponding to each DOE climate zone was selected (listed in Table 3.7) and a simulation using SAM was performed to estimate the total PV energy produced annually for a 5 kW array in each city. Results of the analysis are presented in Figure 3.5. PV arrays located in cities in the southwestern United States, which includes the Hot Dry and Mixed Dry zones, have the most solar PV generation potential. This is due not only to their latitude, but also due to precipitation in the region. For example, Miami, FL is the farthest

Table 3.7: DOE climate zones and selected corresponding cities

US DOE Climate Zone	Location
Mixed Humid	Birmingham, AL
Hot Humid	Miami, FL
Hot Dry	Tucson, AZ
Mixed Dry	Amarillo, TX
Cold	Salt Lake City, UT
Very Cold	Duluth, MN
Marine	Seattle, WA

south of all cities listed, but has a lower annual estimated PV energy generation value than Tucson, AZ or Amarillo, TX due to Miami being located in the Hot Humid climate, where precipitation is high, resulting in greater cloud cover. Performing this analysis in regions with greater PV generation levels would be expected to result in greater cost and energy savings, due to more available PV energy. Regions with lower PV energy generation and higher heating loads, such as Duluth, MN, would have less energy and cost savings.

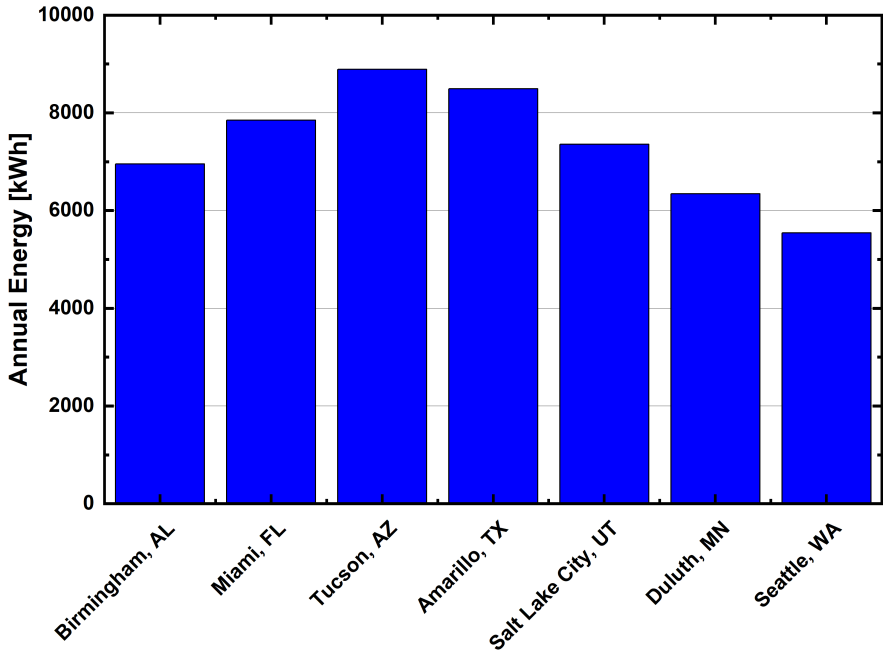

Figure 3.5: Annual solar PV energy production for 5kW array in various geographic locations

In addition to the amount of solar PV energy generated, the structure of the microgrid itself must be considered. In this study, the energy consumption of a single house is optimized. The solar PV generation measured at the neighborhood microgrid is scaled down to simulate PV energy generation at one house and the microgrid is considered to consist of this single-house PV generation, the utility grid connection, and the appliances and other loads within the house. A microgrid that consists of the entire neighborhood and the entirety of the PV array may also be considered. In this case, scheduling of the appliances should be performed in aggregate and the PV energy distributed among houses accordingly. Depending on the behavior of various house occupants, this would result in more or less energy savings for each house in the neighborhood. If loads are successfully scheduled to stagger energy consumption throughout the peak solar PV period, all houses may benefit from shared optimization. However, if one house typically consumes more energy than others using uncontrolled appliances, it would see a greater benefit in sharing the PV array than its neighbors, who might actually see a reduction in energy savings.

The fact that the houses in this neighborhood are completely electric must also be considered. Many homes in the U.S., particularly in colder regions, have gas space heating, water heating, and cooking appliances. The introduction of gas appliances would result in lower cost and energy savings, due to the inability of shifting electrical load for these appliances. Local utility policies and the effect of customer control on thermostat set points must also be considered. In this neighborhood, it is assumed that solar PV fed to the grid is not compensated monetarily. However, in regions with PV buyback policies, the homeowner may benefit by selling more PV energy back to the grid and achieving greater cost reduction. In this neighborhood, the occupants also control the thermostats. The set

points dictate the amount of energy consumed by the HVAC system and water heater. In a setting such as a college dormitory, where set points may be centrally controlled, the total energy savings may increase. Electric utility rates also have an impact on results and may be used by the utility to incentivize customers to participate in demand response programs in which controllable appliances are scheduled and controlled by the utility.

CHAPTER 4. OPTIMIZATION FRAMEWORK

Effectively scheduling and controlling the appliances in a house using a HEMS requires optimizing the scheduled energy consumption of the set of controllable appliances within the house. Data gathered from the site, including temperature and power measurements, are supplied to develop models representing the controllable components of the system. These models are used by the optimization algorithm, along with a pricing schedule provided by the utility and a PV energy forecast, to generate the best possible set of control actions for the controllable appliances. A flow diagram of the overall HEMS is provided in Figure 4.1. In this study, the primary optimization goals are maximizing electrical load shifting for demand response (DR) events, minimizing electrical energy cost for the consumer, and maintaining thermal comfort. The peak demand period is designated

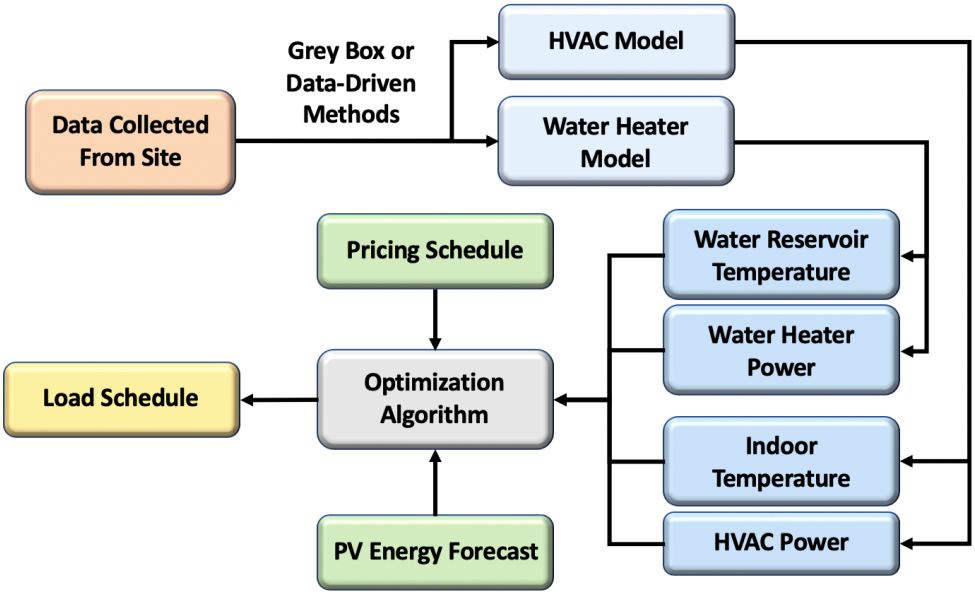

Figure 4.1: Flow diagram of HEMS optimization structure

by the local utility's time-of-use (TOU) rate plan, which establishes the pricing schedule. The pricing schedule uses high prices during peak periods to make operating appliances within the peak period prohibitively expensive. This principle may be used to schedule any type of DR event by varying the pricing schedule. The peak demand period has a higher cost of electricity than other times of the day, thus the minimization of electrical energy consumption during this period also contributes to the minimization of cost for the end user. The optimization problem is developed with the objectives of minimizing cost to the consumer while maintaining a standard of thermal comfort for both hot water and space conditioning. Two methods of optimization are applied to the problem:

1. The problem is formulated as a constrained linear optimization problem and solved with a commercially available solver.
2. The constrained objective function is converted to an unconstrained problem by applying a penalty function rather than constraints and then solved using a novel genetic algorithm.

In the second method, energy consumption scheduling is optimized using both grey-box and black-box models within the fitness function of the genetic algorithm. The grey-box models are based on thermal systems theory but contain data-driven components that help account for noise and measurement error. The black-box models are solely data-driven, although knowledge of thermal systems is used to select the parameters used in modeling. The black-box models must be trained on data from the specific appliance being modeled but have the advantage of being generalized and requiring little or no input from the user. Grey-box and black-box modelling methods are further discussed in Chapter 5.

4.1 Constrained Linear Optimization

To demonstrate the performance of linear optimization problem formulation for the appliance types considered, the optimization problem was formulated in a manner similar to the work of Jin et al. (2017), in which a constrained linear optimization problem was posed and then solved with the commercially available software package CVX. CVX is a modeling software for convex optimization developed for use in MATLAB® (Grant and Boyd, 2013). The SDPT3 solver was used to find the optimal solution using CVX. This approach requires little computational power and would be ideal for deployment to the resource-constrained platforms present in most HEMS systems (Jin et al., 2017). However, the optimization problem in this case is constrained, which may result in infeasible solutions. For example, if the water heater experiences a large-volume water draw when the reservoir temperature is already near the minimum temperature bound, the heat pump may not be capable of reheating the water reservoir to the minimum temperature within the next time step, resulting in an infeasible solution. A sample calculation for this example is provided in Appendix A. The effect of constraints on optimization effectiveness is discussed further in Chapter 6.

The goal of the optimization is to find the set of control variables, U, that minimizes the cost function K as given by Equation (4.1), where U is the set of control variables and x is the initial system state.

$$U_{opt} = \arg\min_{U} K\left(x, U\right) \tag{4.1}$$

The objective function K is formulated as Equation (4.2):

$$K = \sum_{n=1}^{N} C_{\text{n,elec}} P_{\text{n,grid}} \tag{4.2}$$

where $C_{\text{n,elec}}$ is the price of electricity and $P_{\text{n,grid}}$ is the amount of power drawn from the grid at each time n. P_{grid} is calculated using Equation (4.3) through Equation (4.5):

$$P_{\text{grid}} = P_{\text{HVAC}} + P_{\text{WH}} + P_{\text{unctrl}} - P_{\text{PV}} \tag{4.3}$$

$$P_{\text{HVAC}} = U_{\text{HVAC}} P_{\text{nom,HVAC}} \tag{4.4}$$

$$P_{\text{WH}} = U_{\text{WH}} P_{\text{nom,WH}} \tag{4.5}$$

where P_{HVAC}, P_{WH}, P_{unctrl}, and P_{PV} are the values for the amount of power drawn or delivered by the HVAC system, water heater, uncontrolled loads, and PV, respectively. The electricity price vector, C_{elec}, may be thought of as a vector of weights that may be used to prioritize certain objectives of the optimization. Using the pricing vector to prioritize various optimization objectives is discussed further in Chapter 6. Constraints for the objective function (Equations (4.6) through (4.9)) are imposed for the indoor temperature and water heater reservoir temperature. The temperatures of both the indoor conditioned space and water heater reservoir are functions of P and the control variable U.

$$T_{\text{air}}^{\min}(t) \le T_{\text{air}}(t) \le T_{\text{air}}^{\max}(t) \tag{4.6}$$

$$T_{\text{wh}}^{\min}(t) \le T_{\text{wh}}(t) \le T_{\text{wh}}^{\max}(t) \tag{4.7}$$

$$0 \le U_{\text{HVAC}} \le 1 \tag{4.8}$$

$$0 \le U_{\text{WH}} \le 1 \tag{4.9}$$

The constraints maintain the indoor air temperature and water reservoir temperature within a range defined by the user. The models by which the power and temperature values are calculated for each controllable appliance are discussed in further detail in Chapter 5.

4.2 Genetic Algorithm

The problem was then posed as a combinatorial optimization problem that was solved using a genetic algorithm. The optimization problem was formulated to address the issues of non-feasible solutions and solution convergence at the minimum set point found in the constrained linear problem of Section 4.1 and the results are discussed in Chapter 6. In this case, the constraints are converted to a penalty function which is added to Equation (4.2) to yield a new objective function, Equation (4.10). This converts the constrained optimization problem to an unconstrained one.

$$K = \sum_{n=1}^{N} \left(C_{\text{n,elec}} P_{\text{n,grid}} + \Phi_{\text{n}} \right) \tag{4.10}$$

where Φ_n is the penalty function. The penalty function is calculated using Equations (4.11) – (4.13).

$$\Phi_{\text{n}} = \Phi_{\text{n,HVAC}} + \Phi_{\text{n,WH}} \tag{4.11}$$

$$\Phi_{\text{HVAC}}(t) = \begin{cases} T_{\text{air}}(t) > T_{\text{air}}^{\max}(t) & \omega_1 \left(T_{\text{air}}(t) - T_{\text{air}}^{\text{SP}}(t) \right) \\ T_{\text{air}}(t) < T_{\text{air}}^{\min}(t) & \omega_2 \left(T_{\text{air}}^{\text{SP}}(t) - T_{\text{air}}(t) \right) \\ T_{\text{air}}(t) \le T_{\text{air}}^{\max}(t), T_{\text{air}}(t) > T_{\text{air}}^{\text{SP}}(t) & \omega_3 \left(T_{\text{air}}(t) - T_{\text{air}}^{\text{SP}}(t) \right) \\ T_{\text{air}}(t) \ge T_{\text{air}}^{\min}(t), T_{\text{air}}(t) < T_{\text{air}}^{\text{SP}}(t) & \omega_4 \left(T_{\text{air}}^{\text{SP}}(t) - T_{\text{air}}(t) \right) \end{cases} \tag{4.12}$$

$$\Phi_{\mathrm{WH}}(t)=\begin{cases} T_{\mathrm{WH}}(t)>T_{\mathrm{WH}}^{\max}(t) & \omega_1\left(T_{\mathrm{WH}}(t)-T_{\mathrm{WH}}^{\mathrm{SP}}(t)\right) \\ T_{\mathrm{WH}}(t)<T_{\mathrm{WH}}^{\min}(t) & \omega_2\left(T_{\mathrm{WH}}^{\mathrm{SP}}(t)-T_{\mathrm{WH}}(t)\right) \\ T_{\mathrm{WH}}(t)\le T_{\mathrm{WH}}^{\max}(t), T_{\mathrm{WH}}(t)>T_{\mathrm{WH}}^{\mathrm{SP}}(t) & \omega_3\left(T_{\mathrm{WH}}(t)-T_{\mathrm{WH}}^{\mathrm{SP}}(t)\right) \\ T_{\mathrm{WH}}(t)\ge T_{\mathrm{WH}}^{\min}(t), T_{\mathrm{WH}}(t)<T_{\mathrm{WH}}^{\mathrm{SP}}(t) & \omega_4\left(T_{\mathrm{WH}}^{\mathrm{SP}}(t)-T_{\mathrm{WH}}(t)\right) \end{cases} \tag{4.13}$$

where ω_i are weights that may be set to prioritize one outcome over another and may be assigned based on the preference for different comfort parameters. In both the HVAC and water heater models, constraints ω_3 and ω_4 correspond to the penalty for deviating from the temperature set point but not falling outside the temperature dead band. Different weights may be set here to reflect different preferences in deviation from the set point. For example, a user may prefer warmer indoor air temperatures to colder ones and may set greater weight values to penalize temperatures falling below the set point and minimum temperature. Figure 4.2 provides a graphical representation of the regions represented by

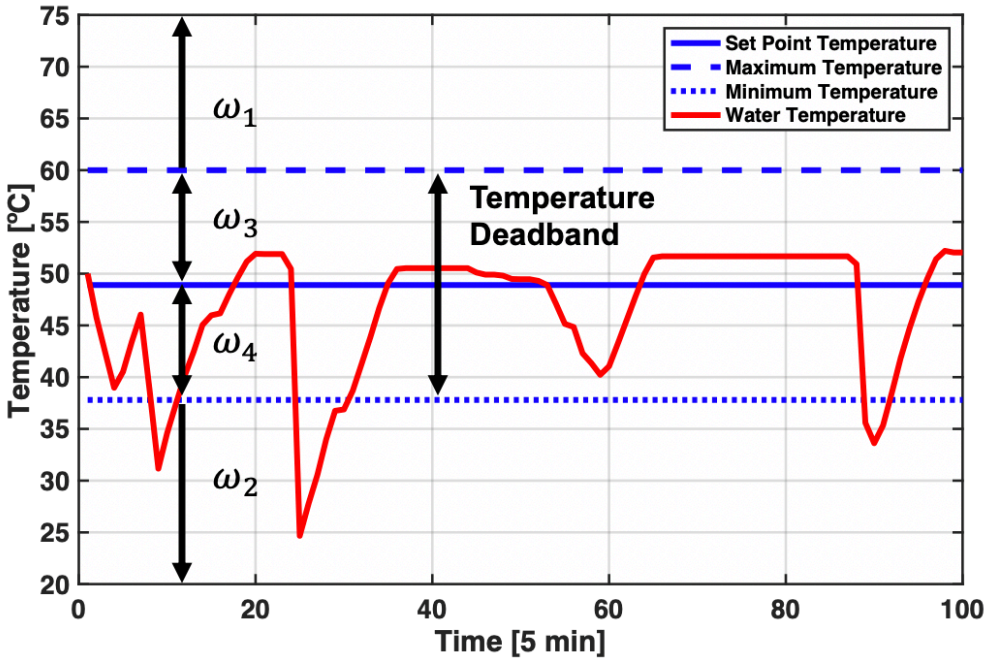

Figure 4.2: Weights for different temperature regions

each weight. In this study, the weights for temperatures falling outside the dead band, ω_1 and ω_2 were determined using a parametric study that evaluated the objective function Pareto front. Use of a penalty function, rather than set of constraints, converts a constrained optimization problem to an unconstrained problem. This allows the objective function to be penalized for producing results that fall outside a prescribed range rather than being confined to a constrained solution space, which in some cases results in infeasible solutions. A sample calculation that illustrates the penalty function effect is included in Appendix A.

4.3 Optimization Solver

The use of machine learning modeling methods in the system model results in a derivative-free optimization problem. It is a combinatorial problem in which an optimal set of binary control variables is selected for 15-minute intervals over an 8-hour time period. Two appliances, the hybrid heat pump water heater and the HVAC system, are included as controllable appliances. In the grey-box modelling case, the three heating components of the water heater (heat pump and two electrical resistance elements) are treated as three binary finite state machines. In the black-box modeling case, the water heater is treated as a single, binary (on/off), finite state machine. Examples of potential solution matrices for the black-box and grey-box optimizations are given in Figure 4.3. This is a permutation problem in which variables may repeat and the order of chosen variables matters; in this case, the number of possible solutions may be calculated using Equation (4.14), where *c* is the number of choices for each variable (two) and *r* is the number of variables. For the grey-box and black-box models, there are 2^{128} and 2^{64} possible solutions, respectively. The number of possible solutions renders an exhaustive search algorithm infeasible.

Possible Solution Matrix: Grey-Box Model

Water Heater Heat Pump	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Water Heater ER Element 1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Water Heater ER Element 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
HVAC System	1	1	1	0	0	0	0	1	1	1	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0

Possible Solution Matrix: Black-Box Model

Water Heater	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
HVAC System	1	1	1	0	0	0	0	1	1	1	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0

Figure 4.3: Optimization solution matrices

$$n = c^{\mathrm{r}} \tag{4.14}$$

Several methods are available for solving derivative-free optimization problems, including particle swarm optimization, simulated annealing, and genetic algorithm (Blum and Roli, 2003). Optimization using a genetic algorithm (GA) solver was chosen for this problem. The schematic of the genetic algorithm is shown in Figure 4.4. The genetic algorithm calculates the fitness over a population of chromosomes (solutions) for each generation and then performs a series of steps to evaluate and alter the population to create a new generation. The algorithm iterates through generations until a stop criterion is reached; this criterion may be a threshold of a parameter that verifies solution population homogeneity or simply reaching the maximum allowable number of generations.

As pictured in Figure 4.4, the execution of the genetic algorithm begins by generating a random set of chromosomes and evaluating the fitness of this population. After the fitness of the initial random population of parent solutions is calculated, a new set of parents must be selected. Parent selection was performed using the tournament selection method (TOS) (Jebari, 2013). In tournament selection, a number of pairs of

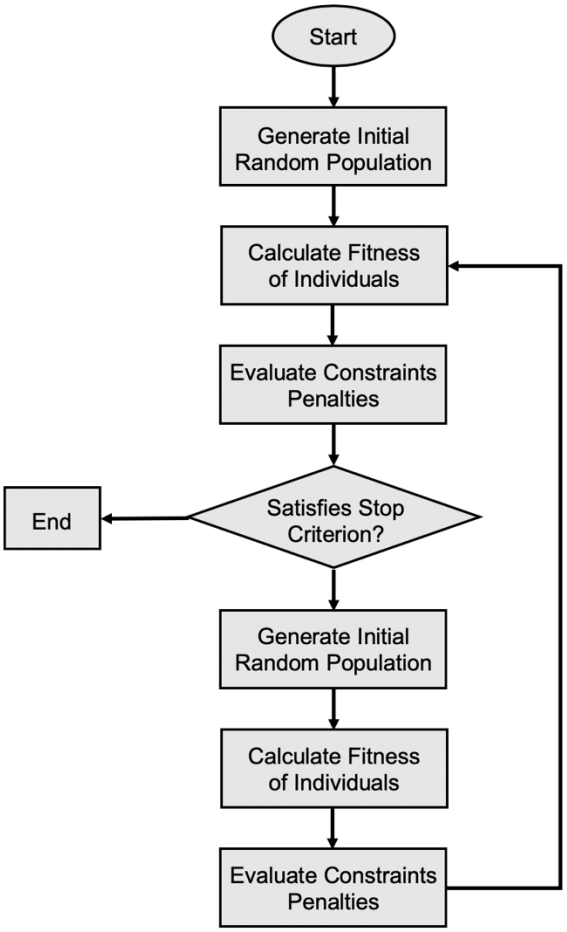

Figure 4.4: Genetic algorithm flow diagram

chromosomes is selected from the population, their fitness functions are calculated, and the better of each pair is chosen to be a parent chromosome for the next generation. From this population, a percentage of elite solutions is preserved for the next generation. The rest undergo crossover and mutation to generate new chromosomes.

The crossover function crosses two parent chromosomes to create two new chromosomes. In this algorithm, multi-point crossover is performed on the chromosome, as pictured in Figure 4.5. The mutation operation then makes an alteration to specific genes

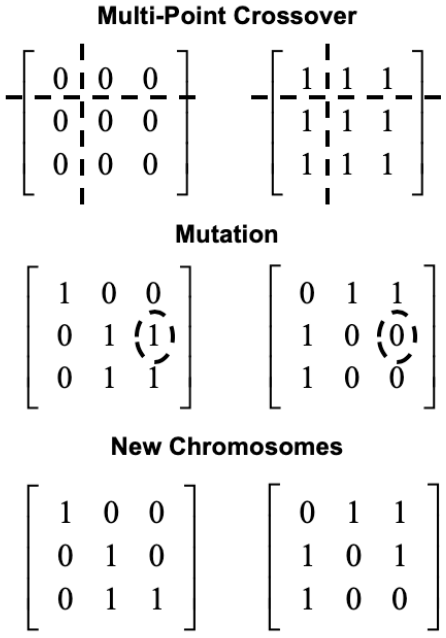

Figure 4.5: Example of crossover and mutation to generate new chromosomes

of the new chromosomes, in this case, switching the binary variable at a random chromosome position. The new solutions and elite solutions are combined in the new population and the process is performed again until reaching the stop criterion. The genetic algorithm described here was written in Python Spyder 3.3.6. The stop criterion for this solver involves inspection of the elite population that is passed to the next generation. In this solver, the solution converges once no new solutions are being preserved in the elite population and passed through to the next generation; e.g., the elite population is homogeneous and the population of loop i is identical to that of loop i+1.

4.3.1 Solver Verification

To test and benchmark the in-house solver, the GA was applied to several common optimization testing functions with known minima. Functions tested included the sphere

function and Rastrigrin's Function (Digalakis and Margaritis, 2001). The global minimum values and search domains used for each function are given in Table 4.1.

Table 4.1: Optimization test functions

Function Name	Global Minimum	Search Domain
Rastrigrin's function	$f(0,0)=0$	$-5 \le x, y \le 5$
Sphere function	$f(0,0)=0$	$-5 \le x, y \le 5$

The sphere function (Equation (4.15)) is a convex function with a known minimum. A plot of the sphere function is provided in Figure 4.6, in which it can be seen that there is an obvious global minimum. This is a straightforward optimization problem, solvable by gradient-based methods, but solved here using the developed genetic algorithm. The solver repeatedly converged on the minimum value of the function, zero, at point (0,0).

$$f(x,y)=x^2+y^2 \tag{4.15}$$

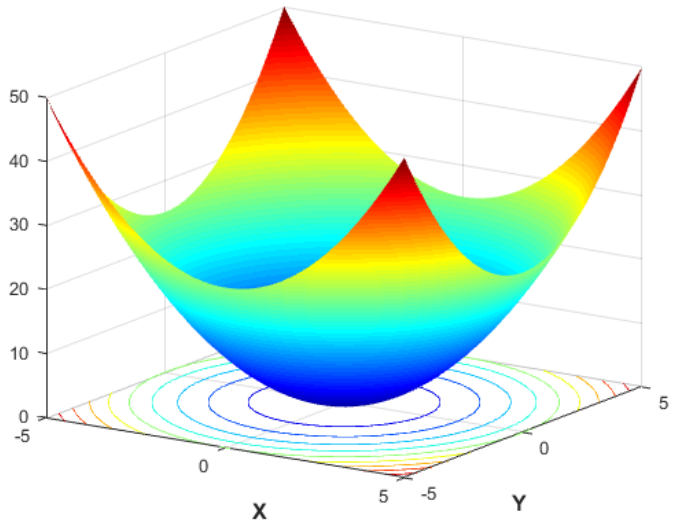

Figure 4.6: Sphere function in three dimensions

Rastrigrin's function, Equation (4.16), is a function commonly used to test genetic algorithm optimization solvers (Digalakis and Margaritis, 2001). A plot of Rastrigrin's function is provided in Figure 4.7. This function is nonconvex, with many local minima and maxima across its grid space. Despite the local minima and maxima, the genetic algorithm reliably converged on the correct solution for Rastrigrin's function.

$$f(x,y) = 20 + x^2 + y^2 - 10\left(\cos(2\pi x) + \cos(2\pi y)\right) \tag{4.16}$$

4.4 Conclusion

Three methods of optimization were developed to address the electrical load scheduling problem encountered by the HEMS. The first method uses a commercial linear solver software package to solve a constrained linear convex optimization problem. The

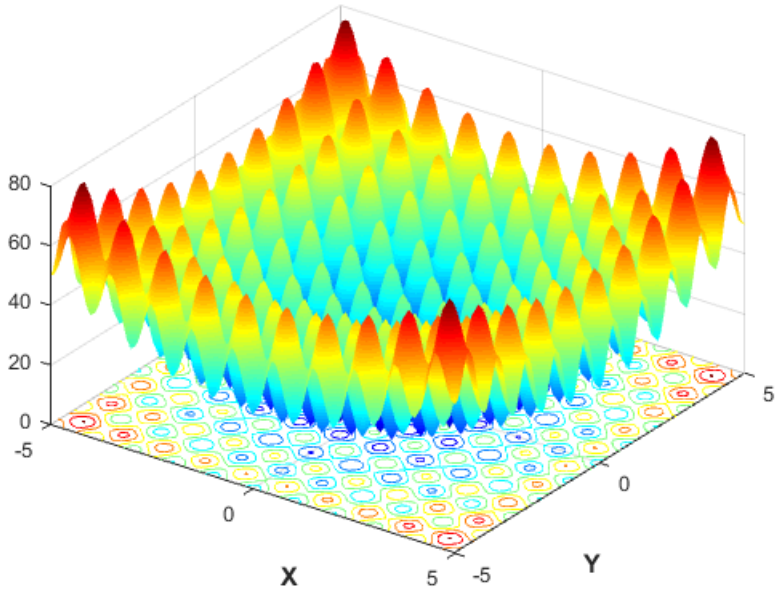

Figure 4.7: Rastrigrin's function in three dimensions

second approach uses a genetic algorithm to solve a linear objective function, with inputs to the objective function found using grey-box modelling methods in one instance and black-box methods in another. The modelling methods for each controllable appliance in the algorithm and the results of the optimization are presented in the following chapters.

CHAPTER 5. THERMAL SYSTEMS MODELING

Systems models are integral parts of each algorithm, providing necessary inputs of power draw and temperature change to the optimization objective function. In this study, both grey-box and black-box models are developed for the controllable components (HVAC and water heater). The grey-box models combine physics-based models with data-driven components that account for noise and measurement error. Grey-box models require inputs like water reservoir volume, system efficiency, house envelope dimensions, system control logic, and house thermal characteristics to successfully model the energy consumption of system components and the temperature response of the environment. With the variety of HVAC and water heater types and a wide range of home construction characteristics, designing a HEMS that offers models specific to every system and site type combination would require hundreds, if not thousands, of model variations. Alternatively, end users may enter the required information into the HEMS, but entering this information is an onerous and error-prone task. For example, to accurately calculate heat transfer from the building envelope, one may be required to enter the total area of windows in the house, which would require accurately calculating the area based on the construction documents for the house or accurately measuring each window in the house. Black-box (data-driven) models offer an advantage over grey-box models by requiring no user input, and instead, requiring only measurements of local transient parameters such as the ambient temperature. In this chapter, the development of both grey-box and black-box models for the water heater and HVAC system is discussed, and the results of the models are compared.

5.1 Heat Pump Water Heater Grey-Box Model

The water heaters used in this study are 0.30 m^3 (80-gallon) electric/heat pump hybrid models. The units have the option of heating water using the heat pump, two 2.5 kW electric resistive elements, or a synchronous combination of electric resistive elements and the heat pump. The heat pump components are located in an enclosure on top of the tank, with the condenser consisting of a helical coil located in the water reservoir. The electric resistive elements are located at the top and bottom of the tank. In a direct load control scenario, the water heater may be turned on and off, but each individual component is not controlled. Therefore, understanding the control logic corresponding to each operational mode is important when developing a predictive energy model for the unit. The available modes of operation provided by the manufacturer are described in Table 5.1. The default operational mode of the water heaters is Energy Saver. In Sparn et al. (2014), a previous model of heat pump water heater by the same manufacturer is evaluated in a laboratory setting and the control logic of the water heater operating in two of the modes available in the current model is described (Table 5.2). These descriptions of control logic

Table 5.1: Water heater operational modes

Mode	Description
Heat Pump	This mode heats with heat pump operation and does not use electric heat during typical heating and demand cycles. This mode has a low recovery but minimizes power consumption.
Electric	Heat is delivered using electric resistance elements. This mode is recommended for use only in maintenance periods and results in maximum power consumption.
Energy Saver	Heat pump and electric heat is optimized, resulting in low power consumption and high recovery. This is the factory default setting.
Vacation	Maintains the tank temperature at approximately 18.3ºC (65°F).
High Demand	Provides the highest recovery while still providing energy savings. The heat pump and electric elements are operated simultaneously.

Table 5.2: Description of water heater modes (Sparn et al., 2014)

Mode	Description
Electric	When the lower heating element thermistor reaches a temperature of 20ºC, both electric resistance elements activate. The lower element is activated before the upper element when a drop in temperature of approximately 0.5ºC is measured by the lower thermistor. The upper element is deactivated once the top thermistor reaches the set point temperature and the lower element remains active until the bottom thermistor reaches the set point temperature.
Energy Saver	When the lower heating element thermistor reads a temperature of 22ºC or less, the heat pump is activated. The heat pump operates alone if the set point is 52ºC or less, but the electric elements are activated, and the heat pump deactivated if the ambient air temperature is outside the operating range stated by the manufacturer. If the set point is at its maximum value of 58ºC, the heat pump and upper element are used in tandem. There is no scenario, other than ambient temperatures outside the heat pump operating conditions, in which the electric resistance elements operate alone.

are for an older model of hybrid heat pump water heater (HPWH), but comparing operational data collected from the current test site with the described control logic shows that the water heater installed in the neighborhood test houses operates following the same control logic when operating in Energy Saver mode, where the operation of the electric resistive and heat pump elements depends on set point temperature and higher set point temperatures result in more frequent activation of the electric resistive elements. The control sequences of several other models of heat pump water heater are also described by Sparn et al. (2014).

In addition to more complex control logic than that of electric resistive water heaters, the heat pump of the hybrid water heater has a coefficient of performance (COP) that varies with operational conditions. The COP is defined as the ratio of thermal energy (Q_{th}) added to the tank to the work input (W_{in}) to the system (Equation (5.1)).

$$COP = \frac{Q_{th}}{W_{in}} \tag{5.1}$$

Past residential energy optimization problems have typically considered electric resistance water heaters (Xu et al., 2014; Jin et al., 2017), which are expected to have an approximately constant COP that approaches unity. The hybrid water heater's heat pump, however, has a COP usually greater than 2.0 that varies with the evaporator (ambient air) and condenser (water reservoir) temperatures. The nominal power value for an electric resistance water heater is constant, while the hybrid water heater must consider the variation of power draw due to heat pump operation and the combined power of both heat pump and electric resistance elements, as well as power draw of the electric resistance elements when they are operating in isolation. The daily COP was calculated for the water heater in heat pump mode for a variety of operating conditions using the method of Sparn et al. (2013). This protocol was developed by the National Renewable Energy Lab (NREL) specifically for field installations of air-source heat pump water heaters. In field installations, the water heater is installed in an occupied house and the parameters available for measurement are typically those outside the water heater that may be obtained without modifying the water heater itself, such as flow rate and water temperature entering and leaving the water heater reservoir. According to this method, COP is calculated using Equation (5.2).

$$COP = \frac{Q_{thermal}}{W_{input}} = \frac{Q_{draws}}{W_{input}} = \frac{\rho_{water} \cdot V_{draws} \cdot c_{p,water} \cdot \Delta T_{draws}}{W_{input} \cdot 3600 \frac{s}{hr}} \tag{5.2}$$

where is the density of water, V_{draws} is the volume of water drawn, $c_{p,water}$ is the specific heat of water, $\Delta T_{draws} = T_{outlet} - T_{inlet}$ is the difference between the outlet and inlet water temperature, W_{input} is the total electrical energy used by the water heater (includes compressor and all auxiliary components) during the day, and $Q_{thermal}$ is the heat removed during water draws.

Figure 5.1 shows the correlation between COP and ambient temperature with varying marker diameters indicating the volume of hot water consumed on a given day. Data from the months of August and January were used to generate the plot; the two distinct point clusters represent the two months. No significant correlation between ambient temperature and COP is seen in this plot. Figure 5.2 shows a stronger correlation between COP and daily flow volume than that seen between COP and ambient temperature in Figure

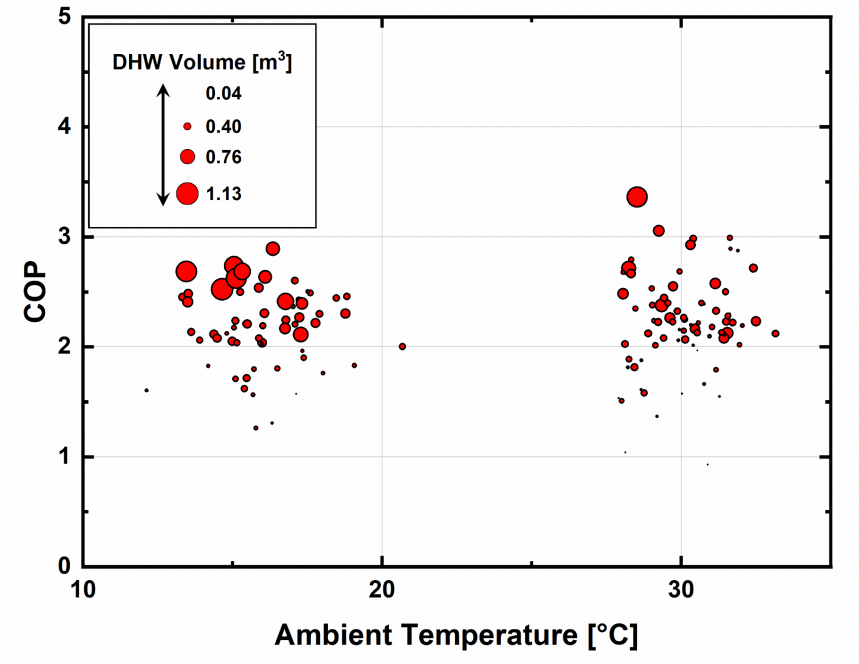

Figure 5.1: Correlation among COP, ambient temperature, and hot water flow volume

5.1. A strong correlation between daily flow volume and COP was also observed in Sparn et al. (2013), but an explanation for this phenomenon was not provided. COP is primarily a function of the difference between heat sink and source temperature, which is reflected in the Carnot COP for heating (Equation (5.3)).

$$COP_{carnot} = \frac{T_H}{T_H - T_C} \tag{5.3}$$

One explanation for the strong correlation with daily flow volume may be that days with higher hot water draw volumes have, consequentially, lower mean water reservoir temperatures, which would reduce the difference between the source and sink temperatures and result in a higher COP. However, Figure 5.3 shows that there does not appear to be a strong correlation between COP and the difference between ambient and reservoir

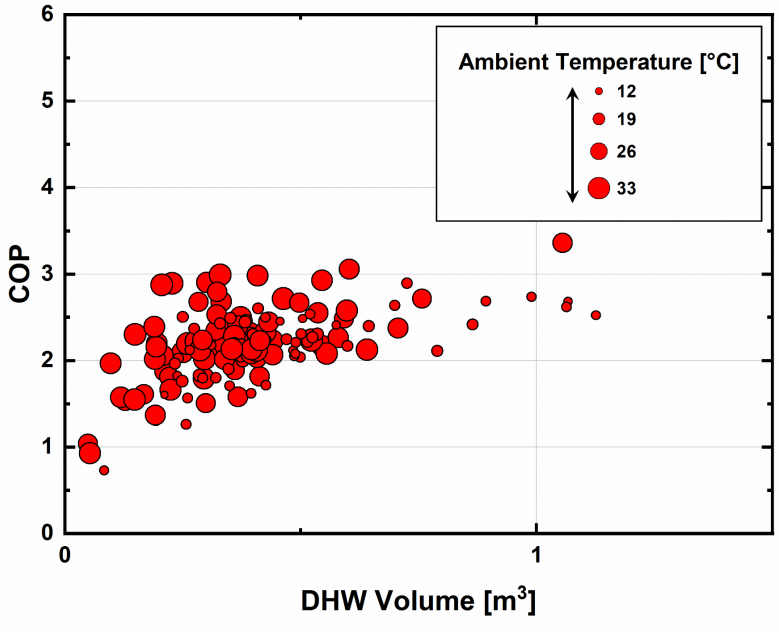

Figure 5.2: Correlation among COP, hot water flow volume, and ambient temperature

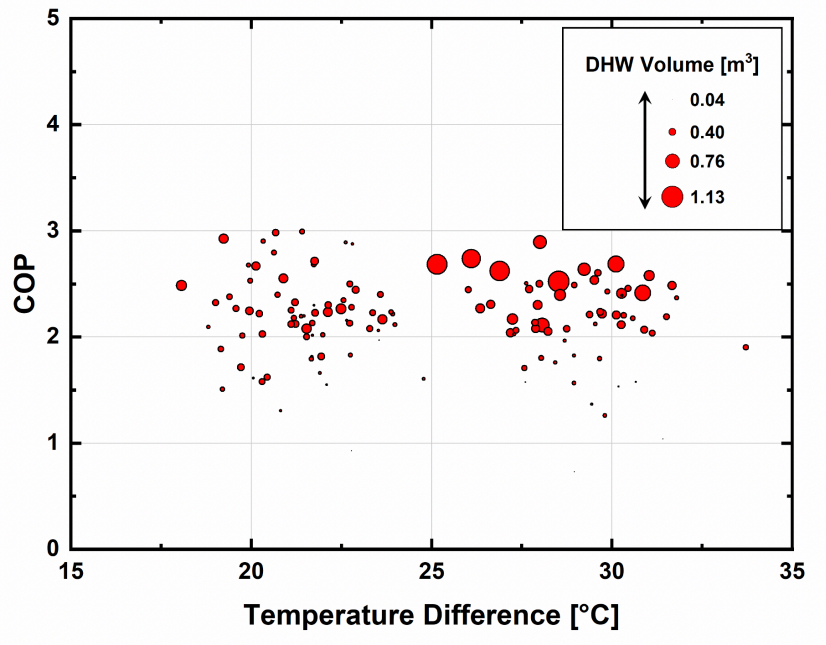

Figure 5.3: Correlation among COP, condenser and evaporator temperatures, and hot water draw volume

temperatures in the field data. To investigate other possible causes for the strong correlation between daily flow volume and COP, a simplified thermodynamic model of the heat pump water heater was developed using Engineering Equation Solver (EES) (Klein, 2018) and is described in detail in the Section 5.1.1. This model was used to calculate COP values for the heat pump at a variety of operating conditions and daily COP for the coupled system of heat pump and water reservoir.

5.1.1 Heat Pump Water Heater EES Model

A model of a hybrid heat pump water heater was developed in EES to generate clean data that were used to understand the correlations among COP, ambient temperature, water reservoir temperature, and hot water flow volume observed in the field data. The data

generated in EES do not have measurement error or the noise that is seen in the field data. This addresses the issue of a correlation being poor due to low-quality data.

The COP of the water heater is calculated on a daily basis, due to the asynchronous nature of the hot water draws and heat pump operation. A quasi-steady state model of a hybrid heat pump water heater with a control scheme similar to that of the water heater installed in the test neighborhood was developed in EES. The heat pump was modeled as a vapor compression system in which the evaporator receives heat from the air surrounding the water heater and the condenser rejects heat to the water reservoir.

A schematic of the cycle is presented in Figure 5.4 and a P-h diagram of the refrigerant side of the cycle is provided in Figure 5.5. From State 2 to State 3, work is done on the refrigerant (R134a) by the compressor to increase the refrigerant pressure (State 4 is the isentropic outlet of the compressor.) From State 3 to State 7, the refrigerant passes through the condenser in which heat is rejected from the refrigerant to the water reservoir;

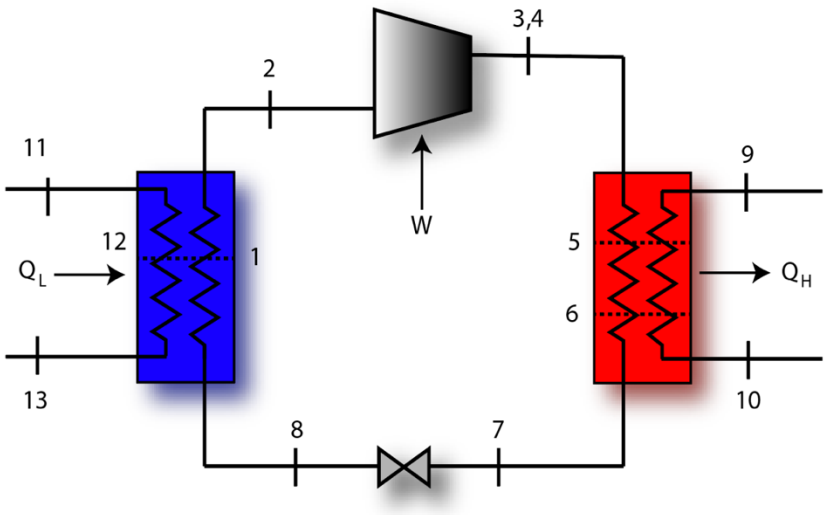

Figure 5.4: Heat pump vapor compression cycle

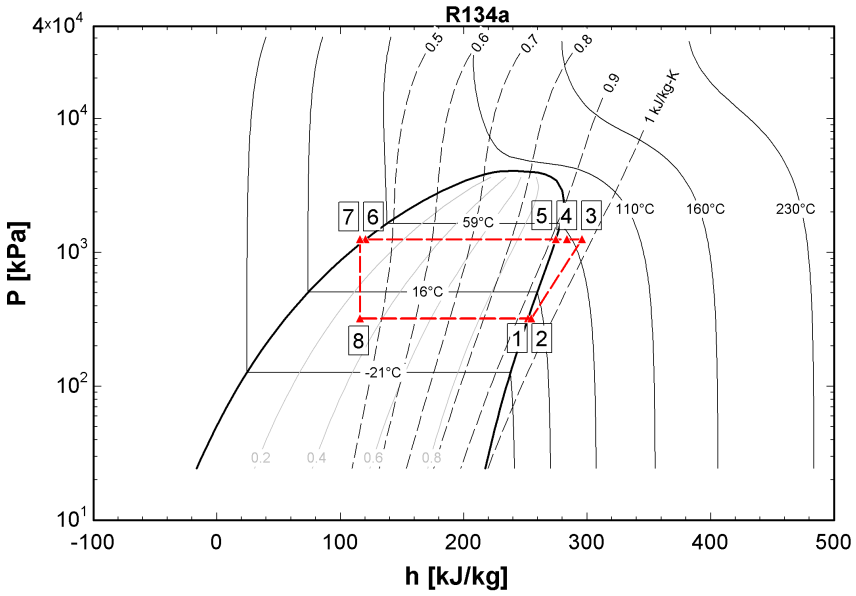

Figure 5.5: P-h diagram of water heater heat pump

States 5 and 6 represent the points at which the refrigerant transitions to a saturated vapor and saturated liquid, respectively. The refrigerant at State 7 is a slightly subcooled liquid. From State 7 to State 8, the refrigerant undergoes isenthalpic expansion and enters the evaporator. Within the evaporator, the refrigerant transitions from a vapor at the inlet quality to a superheated vapor as heat is exchanged between the ambient air passing over the evaporator coil and the refrigerant. The transition from saturated vapor to superheated vapor occurs at State 1, and the refrigerant leaves the evaporator as superheated vapor at a higher temperature at State 2. The refrigerant loop is externally coupled with a water reservoir at the condenser and with ambient air flow at the evaporator. While the heat pump is operating, water in the tank is recirculated using a pump. Thus, the mass flow rate of the pump is taken as the mass flow rate of water across the condenser. State 9 represents subcooled liquid water at a pressure of 65 psia (448.2 kPa) entering the condenser and State

10 represents warmer subcooled water leaving the condenser. Heat transfer from the refrigerant to the water across the condenser is found using Equation (5.4).

$$\dot{Q}_{cond} = \dot{m}_w \left(h_{10} - h_9 \right) \tag{5.4}$$

At State 11, ambient air enters the evaporator. State 12 denotes the point at which the ambient air becomes a saturated vapor. From State 12 to State 13, partial condensation of the humid air takes place and the dehumidified air leaves the evaporator at State 13. Heat transfer at the evaporator is found using Equation (5.5).

$$\dot{Q}_{evap} = \dot{m}_{air} \left(h_{11} - h_{13} \right) \tag{5.5}$$

where $\dot{m}_w$ is the mass flow rate of water across the condenser coil and $\dot{m}_{air}$ is the mass flow rate of air across the evaporator. Initially, the COP of the heat pump was calculated for a variety of ambient temperatures with a fixed water reservoir temperature of 40ºC and plotted in Figure 5.6. The COP increases as the ambient air temperature increases. In Figure

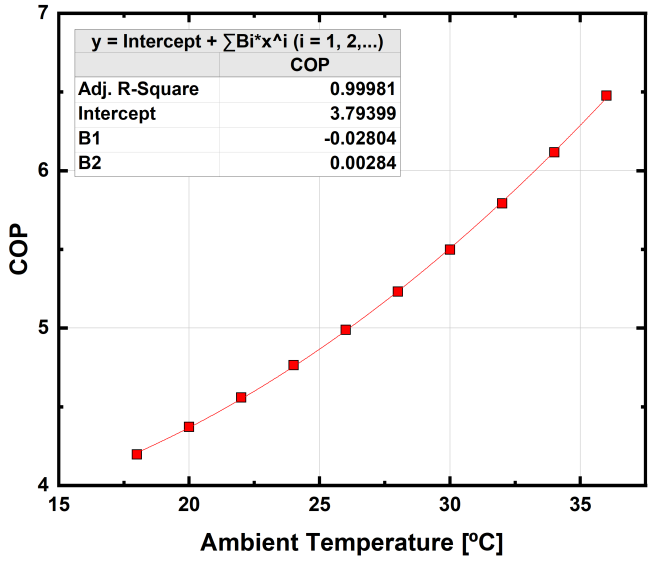

Figure 5.6: Heat pump COP as a function of ambient temperature

5.7, the COP is calculated for a variety of water reservoir temperatures with a fixed ambient air temperature of 20ºC. The COP decreases as the water reservoir temperature increases. Both plots demonstrate that COP is a function of the difference between heat sink and source temperature, and therefore, the pressure ratio of the refrigerant compressor.

The heat pump model was then coupled with a model of the water reservoir, shown in Figure 5.8, to capture the transient effects of water draws. The water reservoir is modeled as a quasi-steady state system, which includes the effect of water draws on system operation. An energy balance on the tank yields Equation (5.6), in which Q_{HP} is the heat added to the water reservoir by the condenser and $Q_{\text{draw,in}}$ and $Q_{\text{draw,out}}$ are the quantities of heat added and removed from the reservoir by hot water leaving the tank and cold water entering. The change in internal energy, ΔU, is calculated using Equation (5.7) and the thermal energy entering and leaving the water reservoir, $Q_{\text{draw,in}}$ and $Q_{\text{draw,out}}$, are calculated using Equation (5.8) and Equation (5.9).

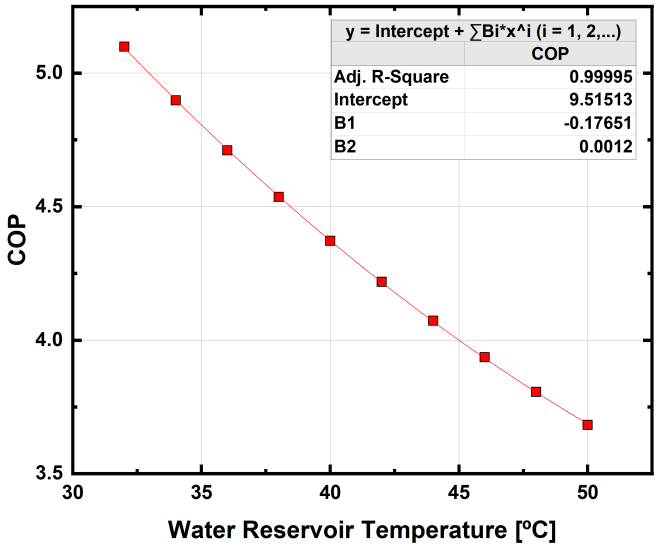

Figure 5.7: Heat pump COP as a function of water reservoir temperature

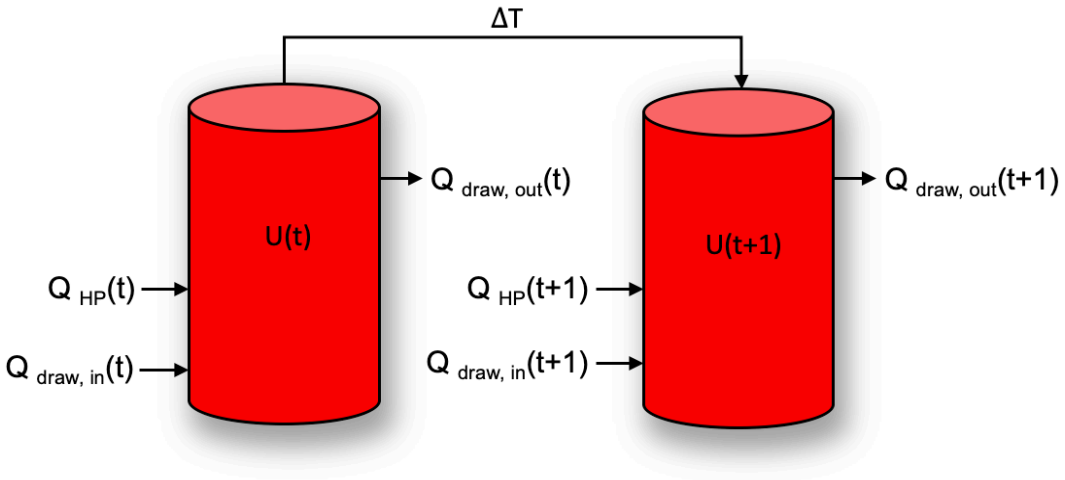

Figure 5.8: Diagram of water reservoir energy balance at two different time steps

$$\Delta U = Q_{draw,in} + Q_{HP} - Q_{draw,out} \tag{5.6}$$

$$\Delta U = m_{w,tank} c_v \left(T_w(t+1) - T_w(t) \right) \tag{5.7}$$

$$Q_{draw,in} = m_w c_p \left(T_{in} - T_{ref} \right) \tag{5.8}$$

$$Q_{draw,out} = m_w c_p \left(T_{tank} - T_{ref} \right) \tag{5.9}$$

where $m_{w,tank}$ is the mass of water in the reservoir and m_w is the mass of water flowing through the tank over a given time interval. One potential source of error when calculating COP using the field data is unquantified heat loss to the ambient during periods in which little water is removed from the tank. For this reason, losses to the ambient were not considered in the EES model, to isolate the heat pump and water flow in and out of the reservoir for analysis. A schedule of water draws was generated using measured volumetric flow data from the neighborhood. Equation (5.2) was used to calculate a daily COP for the water heater. The data generated using this model were then used to calculate the daily

COP over a two-week time period. The resulting COP values are plotted against ambient temperature in Figure 5.9; results show that even though some correlation between ambient temperature and COP is present, the coefficient of determination for this correlation low, with a value of 0.36. This implies a poor correlation between the two variables and poor predictive ability for the resulting correlation. The outlier in Figure 5.9, with a COP value of approximately 11.9, is a result of the water draw and reservoir reheat events falling on different days, as pictured in Figure 5.10; Day 1 is the day on which the outlier COP value of 11.9 was calculated. While several water draw events occurred on Day 1, the water heater did not reach the threshold for reheating until the following day, resulting in a mismatch between the heat removed from the tank and work put into the system. This issue is common, but the result was more pronounced on this particular day. To address this source of error, the time mismatch effect between energy out and work in was removed from the model by developing an artificial water draw data set in which water draws of the

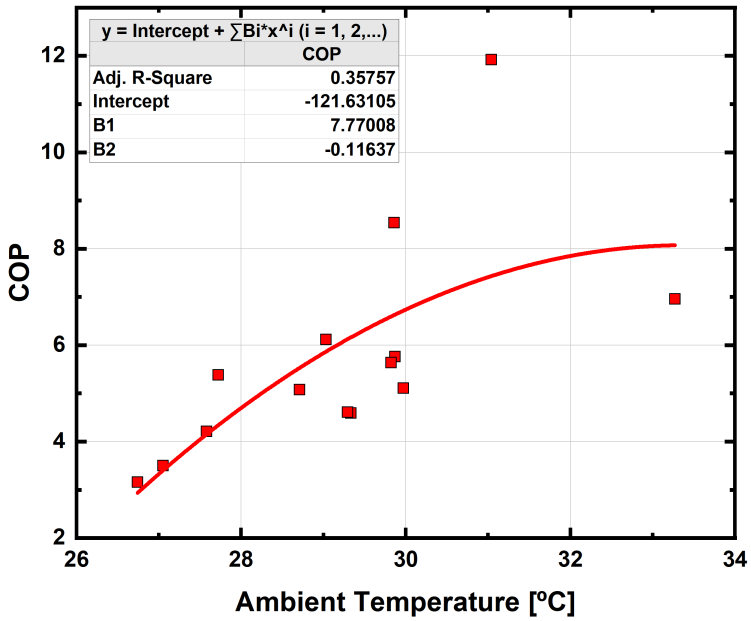

Figure 5.9: EES model daily COP versus daily mean ambient temperature

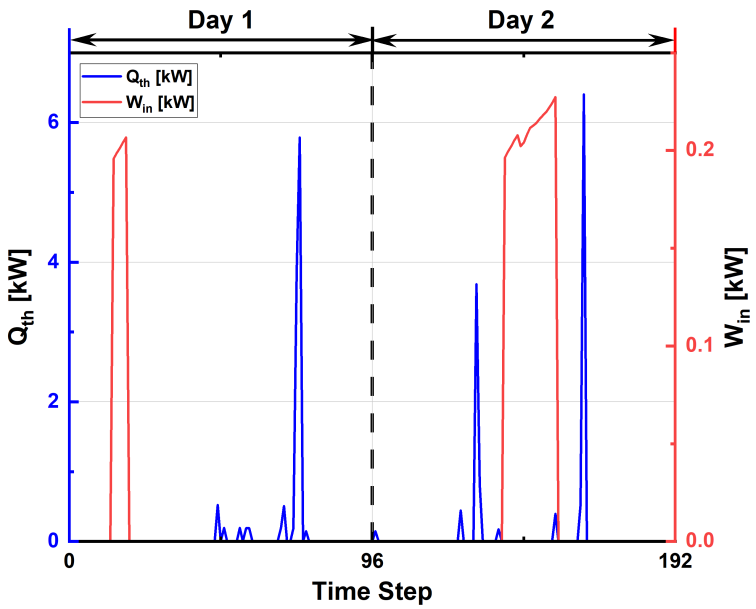

Figure 5.10: Water flow and water reheat

same volume occur early on each day, leaving ample time for the water heater to recover its temperature. Removing the mismatch between water draws and reheat events resulted in the trend between COP and the difference between ambient and water reservoir temperature seen in Figure 5.11. While there appears to be some correlation between the

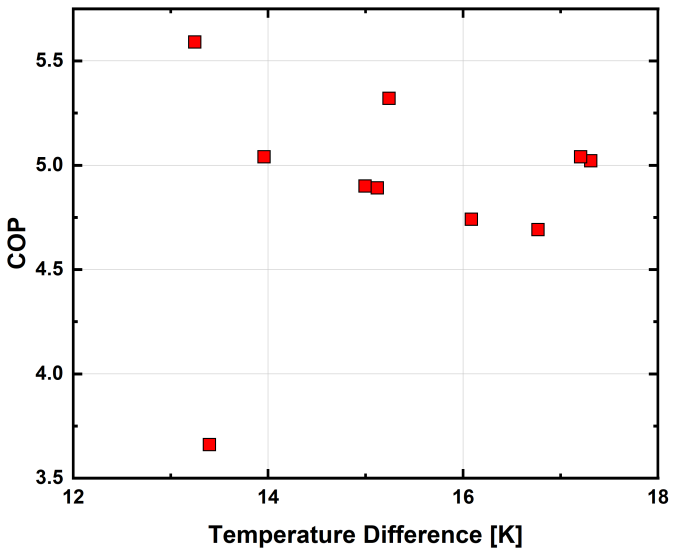

Figure 5.11: Correlation between COP and difference between ambient and water reservoir temperature, calculated using Q_{th}

COP and temperature difference, there are still outliers. Figure 5.12 shows the correlation between COP and temperature difference between ambient temperature and water reservoir temperature calculated using the heat transferred from the heat pump's condenser coil, Q_{HP}. Here, a clear trend is seen between COP and the difference between ambient and water reservoir temperature. The reason for the disparity between the COP calculated using Q_{HP} and the COP calculated using the thermal energy entering and leaving the tank, Q_{th}, lies in the energy storage within the water reservoir. Modeling the heat pump water heater system in EES and then using the clean data generated by the model to develop correlations demonstrated that even in the absence of measurement error and noise, it is challenging to develop a correlation for COP when the thermal energy delivered by the heat pump is determined by calculating the thermal energy entering and leaving the reservoir. Therefore, a constant COP was chosen based on the COP reported by the manufacturer, rather than attempting to develop a correlation from the field data. The COP found in the manufacturer's literature for this water heater was 3.5.

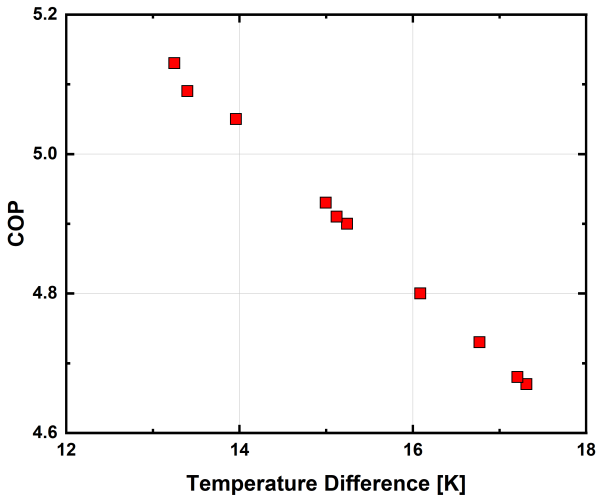

Figure 5.12: Correlation between COP and difference between ambient and water reservoir temperature, calculated using Q_{HP}

5.1.2 Heat Pump Water Heater Grey-Box Accuracy

An energy balance on the water heater reservoir yields Equation (5.10), in which the change of the internal energy of the water reservoir is equal to the net amount of heat transfer due to water flow in and out of the tank plus the heat added by the heat pump (Q_{HP}) and resistive elements (Q_{ER}), less the amount of heat lost to the environment.

$$\Delta U = Q_{draw} - Q_{losses} + Q_{HP} + Q_{ER} \tag{5.10}$$

The change in internal energy of the water reservoir can be expressed as the difference in the average temperatures of the water reservoir at two timesteps, as shown in Equation (5.11)

$$\Delta U = m_{tank} c_v \left(T_w(t+1) - T_w(t) \right) \tag{5.11}$$

The net thermal energy drawn from the tank is a function of the total volume of water drawn over a given time interval τ, and the difference in the average temperature of water entering and leaving the reservoir over the same time interval, as shown in Equation (5.12)

$$Q_{draw} = V_{flow} \rho_w c_p \left(T_{inlet}(t) - T_w(t) \right) \tag{5.12}$$

Thermal energy lost to the ambient is a function of the tank overall heat transfer coefficient, *UA,* which is estimated based on the energy lost from the reservoir due to natural convection, and the average reservoir and ambient temperatures over the time interval. It is calculated using Equation (5.13).

$$Q_{losses} = \tau UA \left(T_w(t) - T_{amb}(t) \right) \tag{5.13}$$

Heat transferred to the water reservoir from the heat pump condenser is calculated using Equation (5.14), where $P_{\text{nom,HP}}$ is the nominal power value for the heat pump, COP is the coefficient of performance determined from the manufacturer's literature, and U_{HP} is a binary control variable. $P_{\text{nom,HP}}$ was determined by finding the value for power that minimizes the mean absolute error between $P_{\text{nom,HP}}$ and the measured power drawn by the water heater over a two-week period.

$$Q_{HP} = \tau P_{nom,HP} COP \cdot U_{HP} \tag{5.14}$$

Heat transferred to the water reservoir from the electric resistive elements is a function of the nominal power of the electric elements, the efficiency of the elements η_{ER} (0.98), and the binary control variable U_{ER}, and is calculated using Equation (5.15),

$$Q_{ER} = \tau P_{nom,ER} \eta_{ER} U_{ER} \tag{5.15}$$

Rearranging Equation (5.11) and substituting into Equation (5.10) gives Equation (5.16), which is used to calculate the predicted mean water reservoir temperature for the next time step, given the conditions and control values of the current time step.

$$T_w(t+1) = T_w(t) + \frac{1}{m_{w,tank} c}\left[Q_{draw} - Q_{loss} + Q_{HP} + Q_{ER}\right] \tag{5.16}$$

Equation (5.16) was then applied to the field data, yielding the accuracy values presented in Table 5.3. Accuracy was evaluated using three metrics: mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2). MAE is evaluated using Equation (5.17) and RMSE is evaluated with Equation (5.18).

Table 5.3: Water heater grey-box model accuracy

Model	RMSE	MAE	R^2
Water Heater Power [W]	149.62	96.14	-
Water Heater Reservoir Temperature [K]	0.383	0.170	0.954

$$MAE = \frac{1}{N}\sum_{i}^{N}\left(y_{pred,i} - y_{meas,i}\right) \tag{5.17}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(y_{pred,i} - y_{meas,i}\right)^2} \tag{5.18}$$

R^2 is the proportion of variance in the dependent variable that is predictable from the independent variable and is calculated using Equation (5.19).

$$R^2 = 1 - \frac{SE_{res}}{SS_{tot}} \tag{5.19}$$

SE_{res} is the residual sum of errors and SS_{tot} is the total sum of squares. These values are found using Equation (5.20) and Equation (5.21).

$$SE_{res} = \sum_{i}\left(y_{pred,i} - y_{meas,i}\right)^2 \tag{5.20}$$

$$SS_{tot} = \sum_{i}\left(y_{meas,i} - \bar{y}_{meas}\right)^2 \tag{5.21}$$

where $\bar{y}_{meas}$ is the mean of the measured data.

The grey-box model was able to predict the mean temperature of the water reservoir for each prediction interval (Equation (5.16)), with a MAE of 0.17ºC. The measured and predicted temperature values are shown in Figure 5.13.

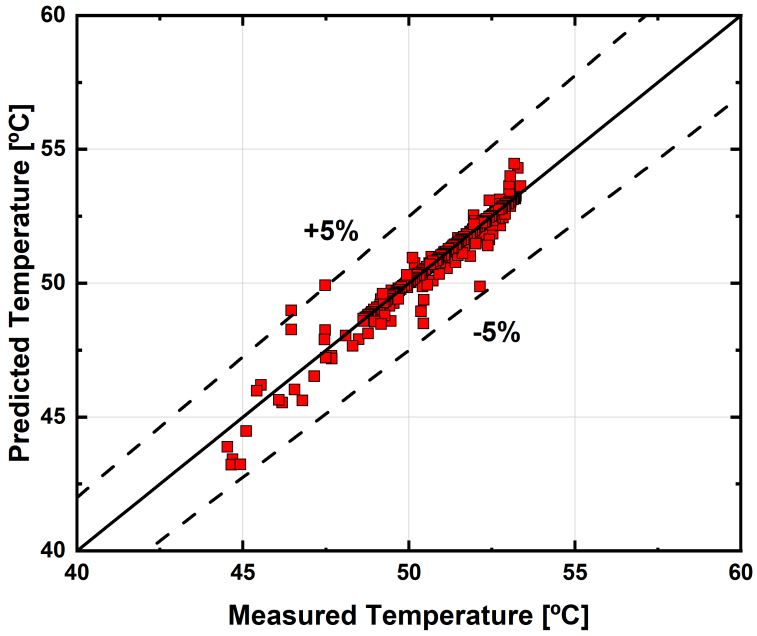

Figure 5.13: Accuracy of grey-box water heater reservoir temperature model

5.2 HVAC Grey-Box Model

The space conditioning needs of the houses are met using a heat pump system with a continuously variable compressor and variable speed fan. The unit has its own proprietary intelligent control system that is expected to remain active during optimization. Therefore, control of individual system components is not possible or desired. Optimization is performed by activating or deactivating the system (ON/OFF). Accurately predicting the system's response to the control action under a given set of external conditions is the requirement for the optimization model. In the house considered in this study, there are two HVAC zones representing the lower and upper floors of the house. The temperature of each zone is measured near the return duct of the respective zone.

Knowing the amount of heat being added or removed from the space by the HVAC unit, a model is then required that captures the change in temperature within each zone. In

addition to the heat transfer between the conditioned spaces and the HVAC system, heat is transferred across the exterior building envelope, through conditioned interzonal boundaries, across interior walls shared with unconditioned spaces such as the attic and garage, and from building occupants and appliances. Energy is also stored in and released from the building's internal mass throughout the day. The interaction of these systems leads to a complex model that is not suitable for use in an optimization algorithm. To account for the interactions between the HVAC system and building envelope, Jin et al. (2017) applied Equation (5.22), in which a linear fit is applied to relevant time-varying parameters that would be included in a more complex model. The coefficients appearing in the equation are intended to account for the effects of static model inputs such as building geometry or insulation R-values.

$$T_{air}^{in}(t+1) = T_{air}^{in}(t) + \gamma_1\left(T_{air}^{out}(t) - T_{air}^{in}(t)\right) + \gamma_2\left(COP_{hvac}^{h} P_{hvac}^{h}(t) + COP_{hvac}^{c} P_{hvac}^{c}(t)\right) + \gamma_3 E_e(t) \tag{5.22}$$

Here, T_{air}^{in} is the temperature within the house, T_{air}^{out} is the outdoor air temperature, COP_{hvac}^{h} and COP_{hvac}^{c} are the COPs of the system in the heating and cooling modes, P_{hvac}^{h} and P_{hvac}^{c} are the nominal power values in the heating and cooling modes, and E_e is the solar irradiance incident on the house. The efficiencies of the system are assumed constant in this approach and are obtained from the manufacturer's equipment manual; in addition, the system has a constant capacity, which allows a nominal power value to be assumed. In contrast, the HVAC systems installed in the test neighborhood have a continuously variable scroll compressor and a variable capacity fan at the indoor air handler. The model of Jin et al. (2017) (Equation (5.22)) predicts the indoor temperature for a single zone house, while

the houses used in the present study have multiple zones. The grey-box model for the HVAC system was modified to account for these differences.

To accommodate the variability of the HVAC system, a separate correlation for power draw was developed. The system is activated when one or more zones deviates from the set point dead band and the capacity of the system varies according to an unknown control logic. As the load on the HVAC system changes, the heat duty of the evaporator and condenser change, affecting the approach temperature of both heat exchangers. Ideally, temperature measurements of both the refrigerant and air side of the system would be available to calculate the approach temperatures and characterize the system performance. However, only indoor return air temperatures and outdoor ambient temperatures are available from the data collected in the field. Considering this, the parameters selected for the power correlation were upstairs and downstairs set point deviations, upstairs and downstairs temperatures, and outdoor temperature.

$$\begin{cases} U_{hvac}=1 \quad P_{hvac}(t)=\gamma_1\left(T_{air}^{Z1}(t)-T_{air}^{SP,Z1}(t)\right)+\gamma_2 T_{air}^{out}(t)+\gamma_3 T_{air}^{Z1}(t)+\gamma_4 T_{air}^{Z2}(t) \\ U_{hvac}=0 \qquad\qquad\qquad\qquad\qquad\qquad P_{hvac}(t)=0 \end{cases} \tag{5.23}$$

The power output of this model was then taken as the power input for the next correlation (Equation (5.24)), which is based on Equation (5.22) but contains a term to reflect the heat transfer between zones. Additionally, power draw of appliances other than the water heater and HVAC were included as P_{unctrl} to represent internal heat gains from appliances.

$$\begin{aligned} T_{air}^{Z1}(t+1) = {} & T_{air}^{Z1}(t)+\gamma_1\left(T_{out}(t)-T_{air}^{Z1}(t)\right)+\gamma_2\left(T_{air}^{Z1}(t)-T_{air}^{Z2}(t)\right) \\ & +\gamma_3 E_e(t)+\gamma_4 P_{unctrl}(t) \end{aligned} \tag{5.24}$$

The measured and predicted power and temperature values with coefficients determined using linear regression are shown in Figures 5.14 and 5.15, respectively. The mean RMSE value across all data sets tested by Jin et al. (2017) was 0.34ºC with a mean R^2 value of 0.92. The temperature difference predicted by Equation (5.24) results in a RMSE of 0.08ºC and an R^2 value of 0.41. This indicates that there is little variance in the output of the model, which results in a low RMSE, but that the predictive ability of the model is relatively poor. Other methods of regression analysis were applied to the problem using the MATLAB® Regression Learner (MathWorks, 2020), including various types of Gaussian process regression and support vector machines. The results of this analysis are presented in Table 5.4, and do not show significant improvement over the linear regression analysis. Possible causes for the poor predictive ability of the model include unmeasured variables such as occupancy levels and infiltration (may fluctuate with open windows or doors), which

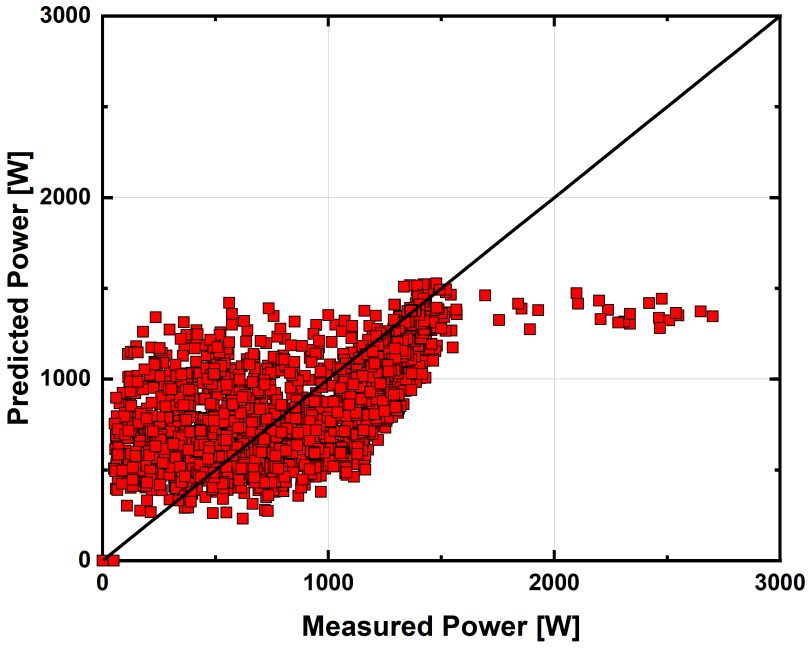

Figure 5.14: Accuracy of grey-box model for HVAC power

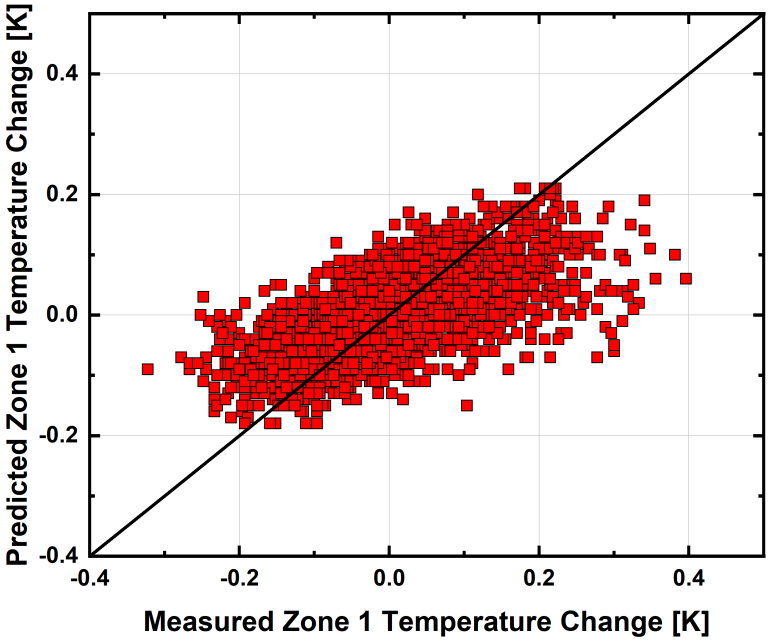

Figure 5.15: Accuracy of grey-box model for HVAC temperature change

influence internal heat gain, and misrepresentation of the actual temperature within the zone. Because temperature measurements are collected at one location, fluctuations in temperature in different areas within the zone may affect the measured temperature in unknown ways. For example, if a space heater was operating discontinuously near the temperature sensor, the change in temperature would become more unpredictable. Additionally, energy transfer between zones introduces additional complexity to the

Table 5.4: MATLAB regression analysis using measured HVAC data

Type	Sub-Type	RMSE [K]	R^2
Linear Regression	Linear	0.084	0.41
	Robust Linear	0.084	0.40
Gaussian Process Regression	Rational Quadratic	0.078	0.48
	Squared Exponential	0.080	0.47
	Matern 5/2	0.079	0.48
	Exponential	0.078	0.49
Support Vector Machine	Linear	0.085	0.40
	Quadratic	0.083	0.42
	Cubic	0.084	0.41

system. To investigate the source of error in the model, a simplified thermodynamic model of the HVAC system and house was developed in EES. The same analyses were then applied to the EES-generated data to determine the potential accuracy of the grey-box and black-box modeling approaches. The EES models are described in Section 5.2.1.

5.2.1 HVAC EES Model

A vapor compression HVAC system (depicted in Figure 5.16) serving a single zone was modeled with a 10 kW cooling capacity. The system model operates in a manner similar to the system described in Section 5.1.1, with the only difference being that the condenser in this case is air-coupled. The HVAC system model was coupled with a model of the building envelope developed following the method of Cui et al., 2019 and outlined in Equations (5.25) through Equation (5.28). Figure 5.17 represents the RC network developed to model the thermal characteristics of the house.

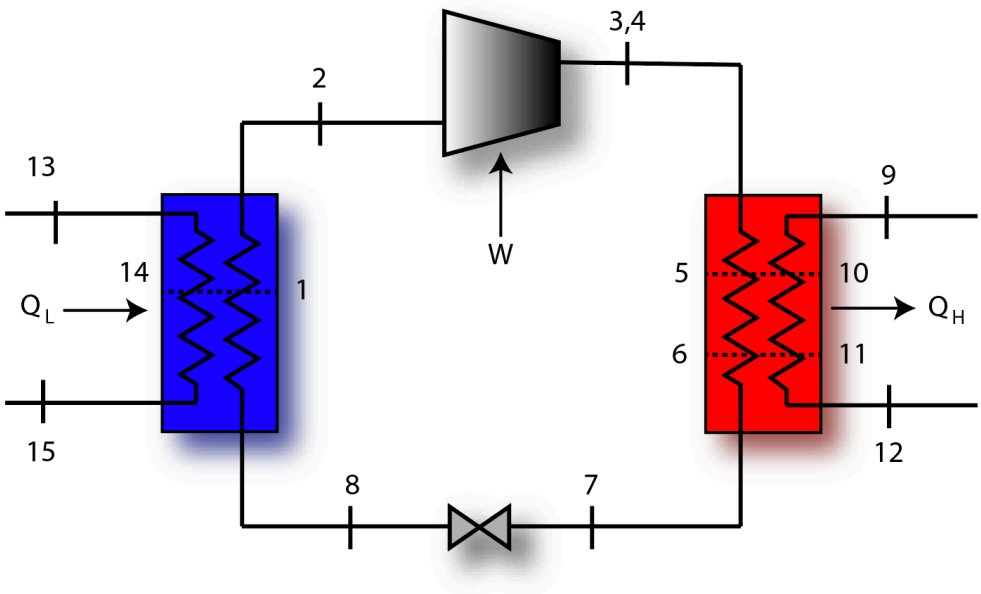

Figure 5.16: EES HVAC model flow diagram

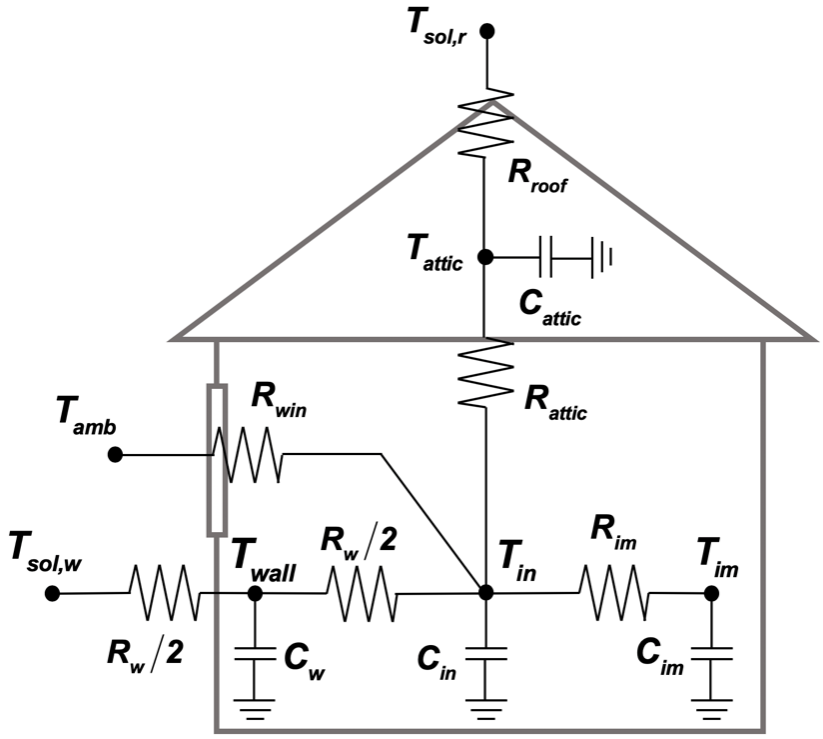

Figure 5.17: HVAC RC model diagram

$$C_w \frac{dT_{wall}}{dt} = \frac{T_{sol,w}(t) - T_{wall}(t)}{R_w/2} + \frac{T_{wall}(t) - T_{in}(t)}{R_w/2} \tag{5.25}$$

$$C_{in} \frac{dT_{in}(t)}{dt} = \frac{T_{wall}(t) - T_{in}(t)}{R_w/2} + \frac{T_{attic}(t) - T_{in}(t)}{R_{attic}} + \frac{T_{im}(t) - T_{in}(t)}{R_{im}} + Q_{IHG} + Q_{HVAC} + Q_{solar} \tag{5.26}$$

$$C_{attic} \frac{dT_{attic}(t)}{dt} = \frac{T_{sol,r}(t) - T_{attic}(t)}{R_{roof}} - \frac{T_{attic}(t) - T_{in}(t)}{R_{attic}} \tag{5.27}$$

$$C_{im} \frac{dT_{im}(t)}{dt} = -\frac{T_{im}(t) - T_{in}(t)}{R_{im}} + Q_{IHL} + Q_{AC} + Q_{solar} \tag{5.28}$$

where $T_{sol,w}$ is the surface temperature of the external wall, $T_{sol,r}$ is the external temperature of the roof, T_{wall} is the surface temperature of the interior side of the envelope, T_{in} is the temperature of air within the house, T_{attic} is the temperature of air in the attic, and T_{im} is the lumped temperature of the internal mass of the house. Q_{IHL} is the internal heat load, Q_{AC} is the heat transferred by the HVAC system, and Q_{solar} is the irradiance incident upon the surface of the house. The resistance R and capacitance C values were determined using information regarding insulation and geometry from the house construction documents and are given in Table 5.5.

Table 5.5: Resistance and capacitance values for house energy model

	Resistance R [K W^{-1}]	**Capacitance C [J K^{-1}]**
External Wall	0.008	9.7×10^6
Attic	0.050	-
Roof	0.003	-
Window	0.006	-
Internal mass	0.007	2.0×10^7
Internal air	-	8.2×10^5
Calculated from house dimensions and material thermal properties provided in the construction documents.		

The same method of multivariate linear regression used to find the coefficients of Equation (5.23) and Equation (5.24) was applied to the data generated by the EES model and the resulting measured and predicted temperature change and power are plotted in Figure 5.18 and Figure 5.19, respectively. The accuracy of the resulting models is given in Table 5.6. The results of applying linear regression to the simulated data demonstrate that with a clean dataset for a single zone, an accurate model may be developed.

Table 5.6: Linear regression accuracy for HVAC EES data

Model	**MAE**	**RMSE**	**R^2**
HVAC Power [W]	1.90	2.45	0.998
Indoor Temperature Change [K]	0.063	0.083	0.980

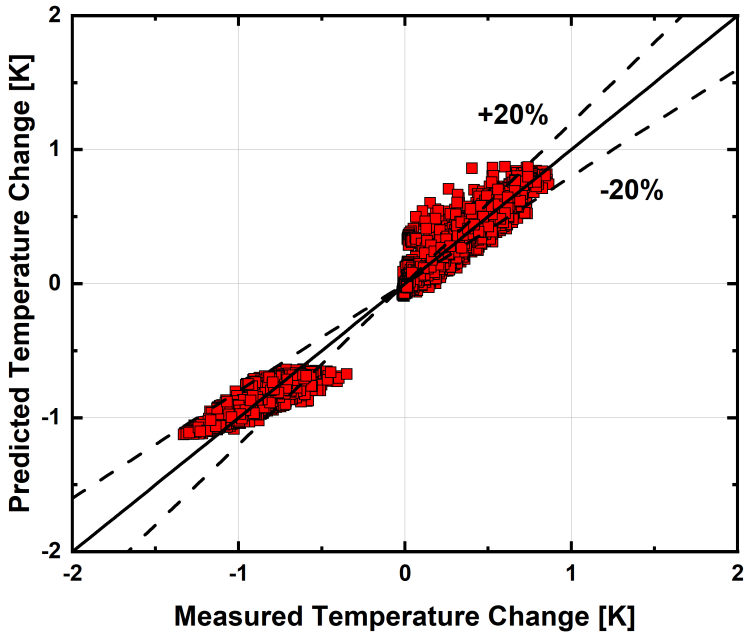

Figure 5.18: Accuracy of grey-box indoor temperature model using EES data

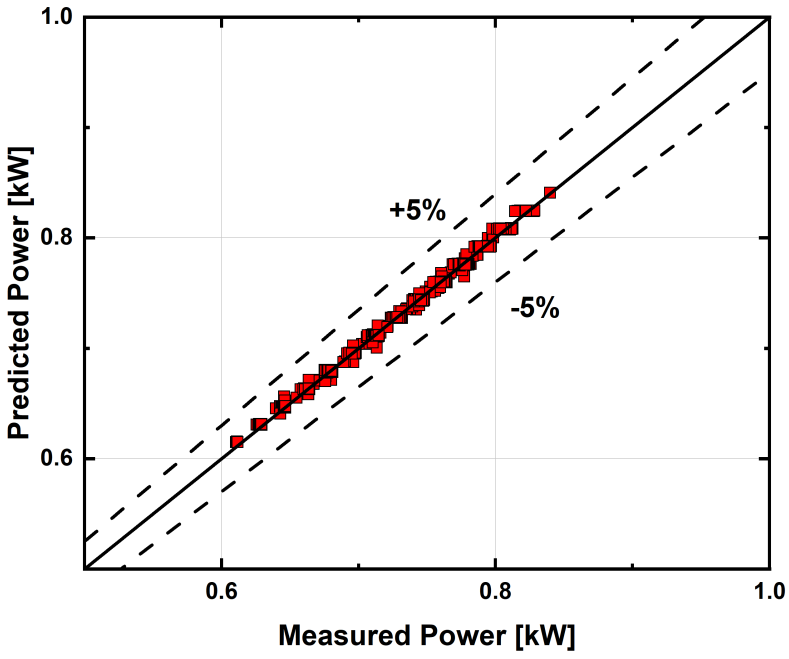

Figure 5.19: Accuracy of HVAC power grey-box model using EES-generated data

5.3 Data-Driven Modeling

Many of the black-box modeling methods used today are decades old but have only recently gained attention due to the increased processing power of computers and the abundance of available data. Black-box models are used in this study to approximate the function that will be used for the space conditioning and water heating models within the optimization algorithm. Extreme gradient boosting was chosen for this application.

5.3.1 Extreme Gradient Boosting (XGB)

Extreme gradient boosting (XGB) is a variant of gradient tree boosting algorithm. XGB produces a predictive model using an ensemble of weaker prediction models, typically decision trees, and is suitable for both classification and regression problems (Chen, 2016). XGB has been reported to produce more accurate results than familiar machine learning algorithms like artificial neural networks and support vector machines when applied to residential energy modeling problems (Cui et al., 2019). The particular XGB algorithm used in this study is XGBoost, developed by Chen (2016). Each model was trained using two weeks of data.

5.3.2 Water Heater Model

Black-box models of the water heater are desirable because they may adapt to any tank volume, mechanical specification, or control sequence, if properly trained. The black-box models for the water heater were developed using XGB and comparisons were made between the black-box model and the grey-box model described in Section 5.1. Several metrics were used for comparison, including the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). The black-box model for the

water heater follows the path shown in Figure 5.20. Inputs identified earlier in the grey-box modeling process are supplied to the first layer of XGB algorithm to predict the power draw of the water heater for a given period. The output of this model (power) is then provided as an input, along with hot water mass flow rate, ambient temperature, and operational mode, to the second XGB model, which predicts a water reservoir temperature change and updates the water temperature for the next iteration. One important advantage of this method of modeling is the lack of a need for detailed knowledge of the system's control logic. Capturing the behavior of the system during the training period allows the algorithm to infer a control logic. The grey-box water heater modeling method assumed knowledge of the control logic as well as the ability to control individual system components, which would not be practical if the model was used in a demand response-type scenario where a wide variety of water heater model types must be controlled. The resulting model accuracy metrics for XGB applied to field water heater data with a two-week training period are provided in Table 5.7. Correlations between the measured and predicted power data are provided in Figure 5.21 and Figure 5.22.

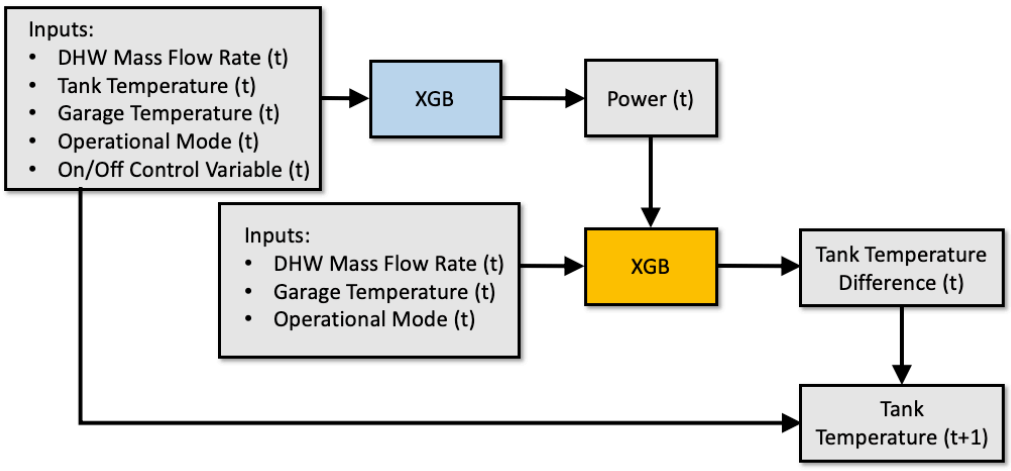

Figure 5.20: Water heater black-box model flow diagram

Table 5.7: Water heater black-box model error

Model	RMSE	MAE	R^2
Water Heater Power [kW]	0.018	0.013	0.973
Water Heater Reservoir Temperature [K]	0.897	0.295	0.900

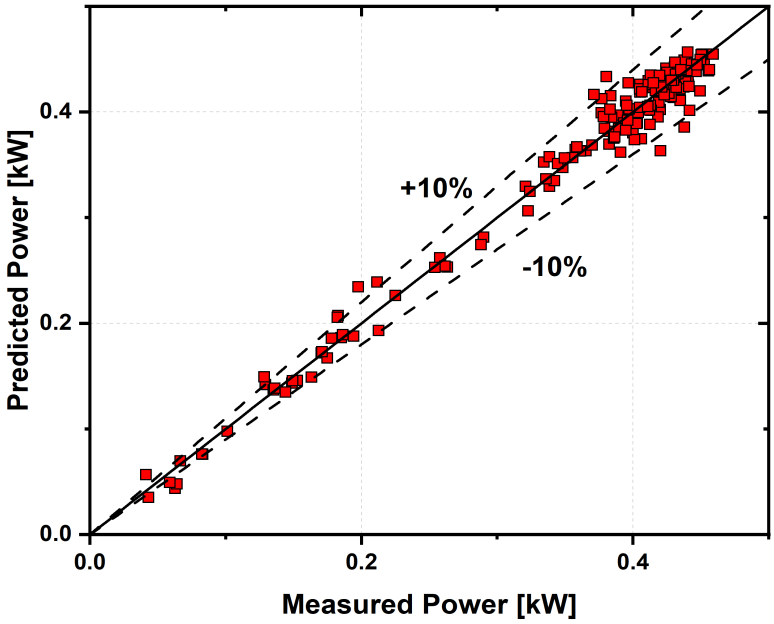

Figure 5.21: Accuracy of water heater power black-box prediction

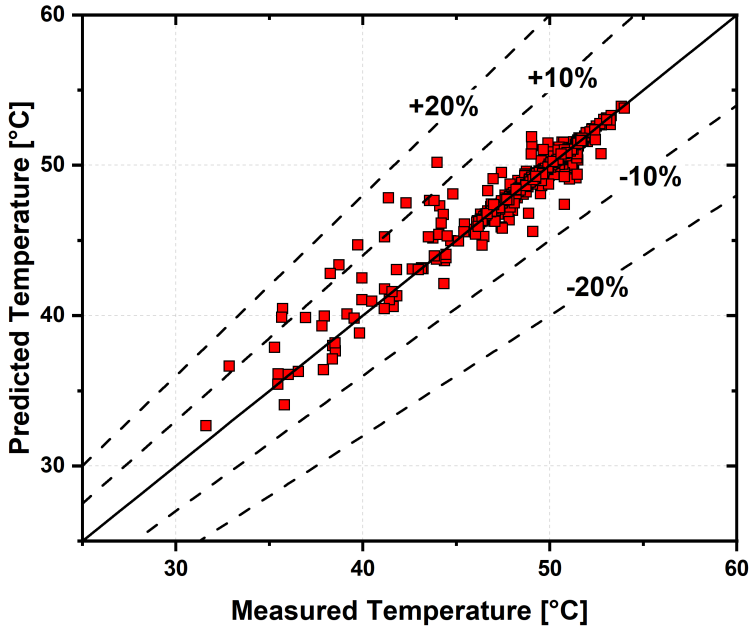

Figure 5.22: Accuracy of water heater temperature black-box prediction

5.3.3 HVAC Model

The method of modeling the HVAC system was similar to that of the water heater, with relevant input parameters determined from the insights gained in the grey-box modeling phase. This model follows the steps shown in Figure 5.23. The deviations from the temperature set point, current zone temperature, irradiance incident on the house, outdoor temperature, operational mode, and the ON/OFF control variable are used as inputs to determine the power consumed by the HVAC unit at a given time. This power, along with the previously mentioned inputs, are then supplied to the second XGB algorithm to predict the change in zone temperature. Models were trained on the data gathered in the field as well as the data generated using the HVAC EES model. Correlations between the predicted and measured power and predicted and measured indoor temperature change are shown in Figure 5.24 and Figure 5.25. The model accuracy for the field data and EES-generated data are given in Table 5.8 and Table 5.9, respectively. Comparing the data

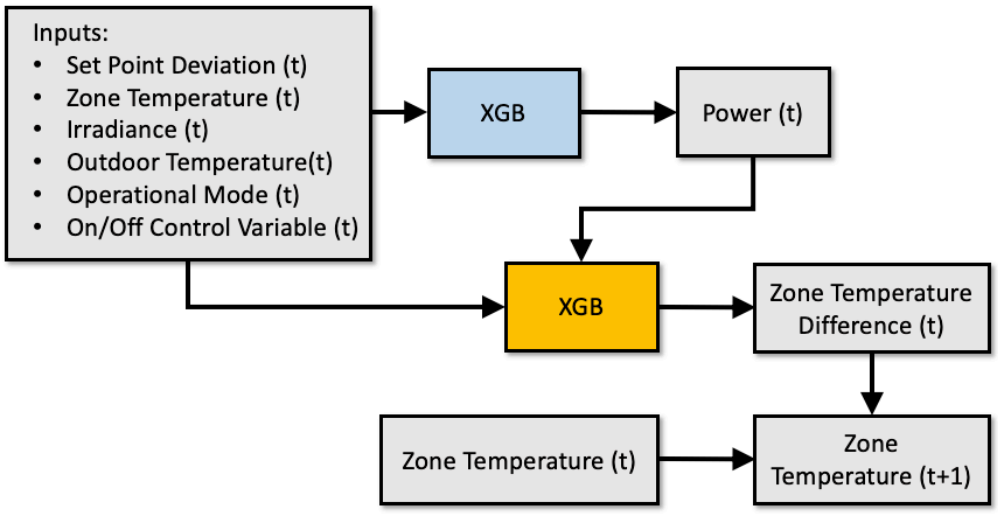

Figure 5.23: HVAC black-box model flow diagram

presented in Table 5.8 and Table 5.9, it is concluded that the black-box model of the HVAC system using field data was worse than the performance of the grey-box HVAC model.

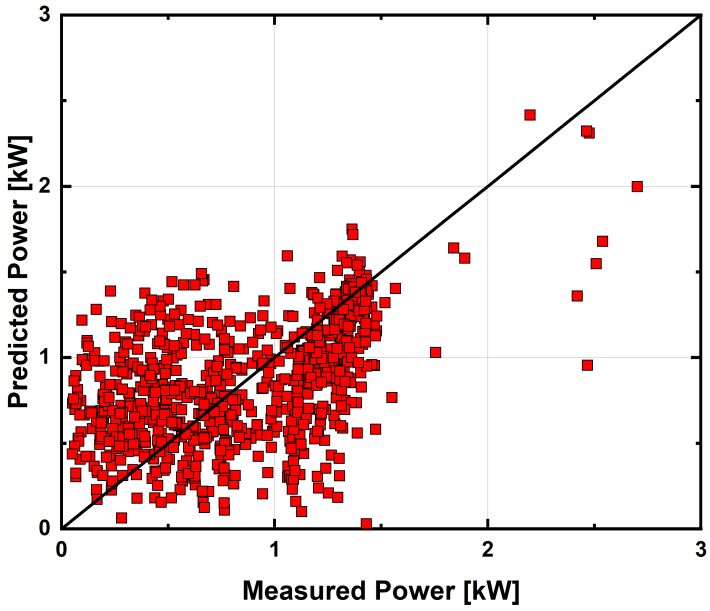

Figure 5.24: Accuracy of HVAC power black-box prediction using measured site data as input

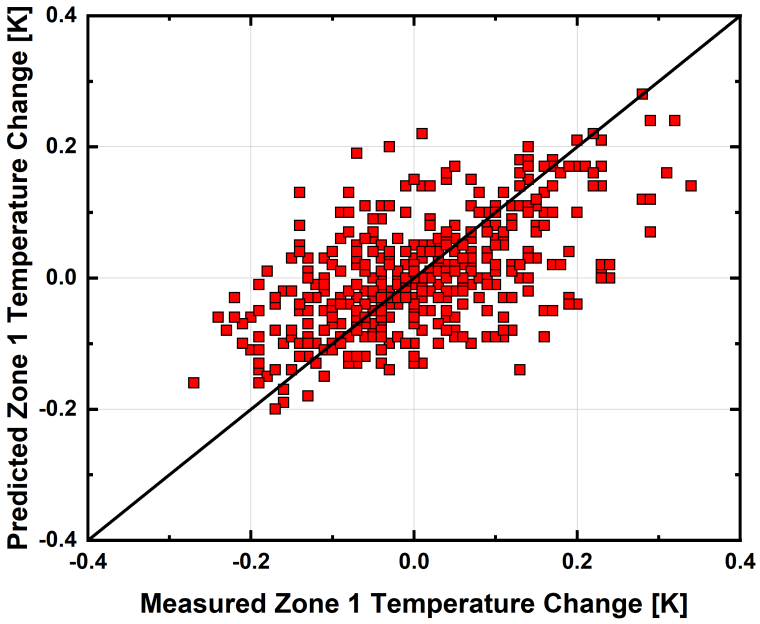

Figure 5.25: Accuracy of indoor temperature black-box prediction using measured site data

Table 5.8: HVAC black-box model accuracy (measured data)

Model	RMSE	MAE	R^2
HVAC Power [kW]	0.423	0.328	0.149
Zone 1 Air Temperature [ºC]	0.093	0.071	0.300
Zone 2 Air Temperature [ºC]	0.143	0.096	0.228

The power prediction using the measured field data had an MAE of 0.328 kW and RMSE of 0.423, while the power prediction using the EES data had an MAE of 0.001 kW and RMSE of 0.001 kW. The indoor temperature was also predicted more accurately using the EES data, with an MAE of 0.042ºC and RMSE of 0.066ºC, as compared with MAE values of 0.071ºC and 0.096ºC for Zone 1 and Zone 2, respectively, and RMSE values of 0.093ºC and 0.143ºC for Zone 1 and Zone 2, respectively, in the model using field data. There are a variety of sources of error for the HVAC models developed using field data, including the lack of information regarding internal heat gain, uncertain or constant sensor measurements, and incomplete information regarding interzonal heat transfer.

The black-box models for predicted power and indoor temperature difference using the EES-generated data were much more accurate than the models developed using measured field data, which may be seen by comparing the accuracy plots of the black-box models developed with field data (Figure 5.24, Figure 5.25) to the accuracy of the models developed using the EES-generated data (Figure 5.26, Figure 5.27). When comparing the results of the black-box and grey-box modeling analyses using EES-generated HVAC data, the black-box models were more accurate than the grey-box models, with the MAE for the power prediction using the black-box model 50% lower than that of the grey-box model

Table 5.9: HVAC black-box model accuracy (EES-generated data)

Model	RMSE	MAE	R^2
HVAC Power [kW]	0.001	0.001	0.999
Indoor Air Temperature [ºC]	0.066	0.042	0.988

and the MAE for the black-box model of temperature change 33% lower than that of the grey-box model.

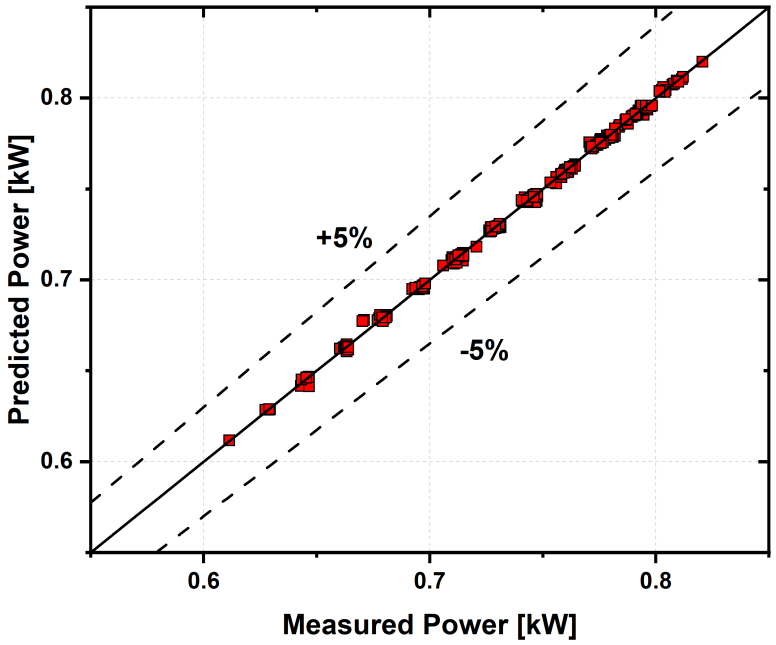

Figure 5.26: Accuracy of HVAC power black-box modeling method applied to EES-generated data

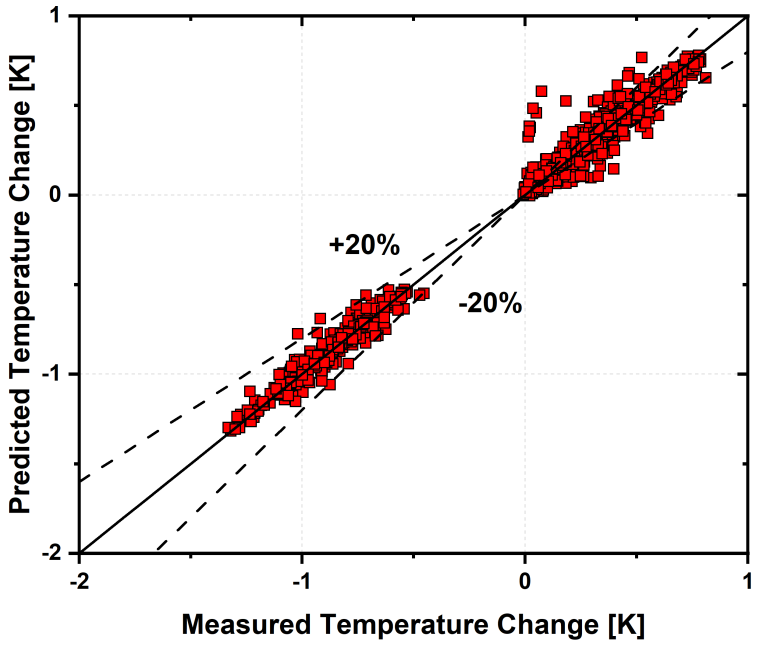

Figure 5.27: Accuracy of indoor temperature change black-box modeling method applied to EES-generated data

5.4 Model Comparison

Both grey-box and black-box models were developed for the HVAC and water heater. Models for the HVAC system were developed using data collected at the test site, as well as data generated using a simplified EES model. Two measures of error, root mean square error (RMSE) and mean absolute error (MAE) are used for comparing the black-box and grey-box models. The method for calculating these errors was shown in Section 5.1.2, Equation (5.17) and Equation (5.18). The error was calculated for each model with respect to the baseline data that it was attempting to predict. Error for the HVAC system power draw was low due to low variability in the power draw when the system was operating and the fact that no noise was present, because the HVAC data were generated using a simulation. The water heater power had higher error for power draw, due to a nominal power value being selected in the grey-box model that did not account for changes in the water heater power draw due to fluctuations in heat exchanger duty. The black-box model was more accurate, with errors less than 20 W. As relative errors, the HVAC power draw error would be near zero, but the water heater would be approximately 37% in the grey-box case and 5% in the black-box case. Better model performance in the HVAC models may be attributed to these models being developed using simulated data without measurement uncertainty and less variability than real-world data gathered in the field. Relative error for the indoor air temperature prediction is negligible and relative error for the water heater temperature prediction is 0.3-2.0%. The difference in these relative errors may also be attributed to the difference in developing the models using simulated data as compared with real-world data. The black-box models outperformed the grey-box models when modeling both the power draw and resulting indoor temperature of the HVAC system, as shown in Figures 5.28 and 5.29, respectively. This result may be attributed to

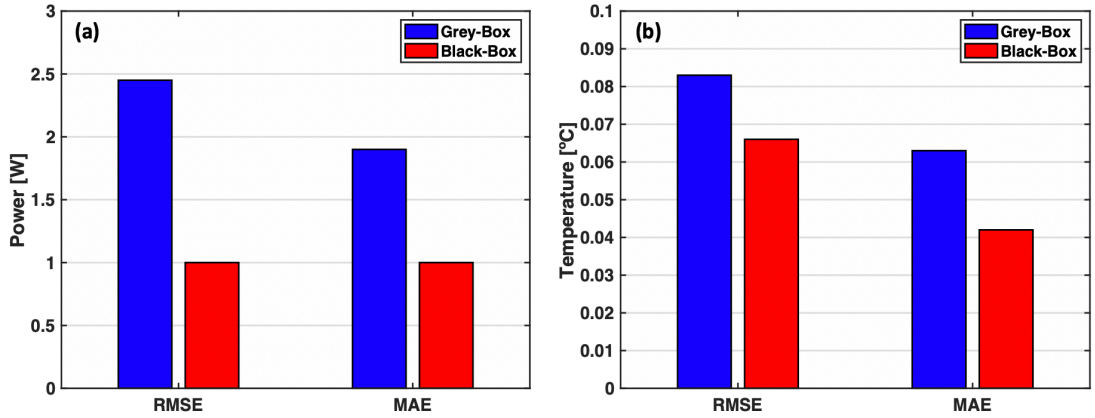

Figure 5.28: Comparison of grey-box and black-box modeling accuracy for (a) power and (b) temperature of HVAC system

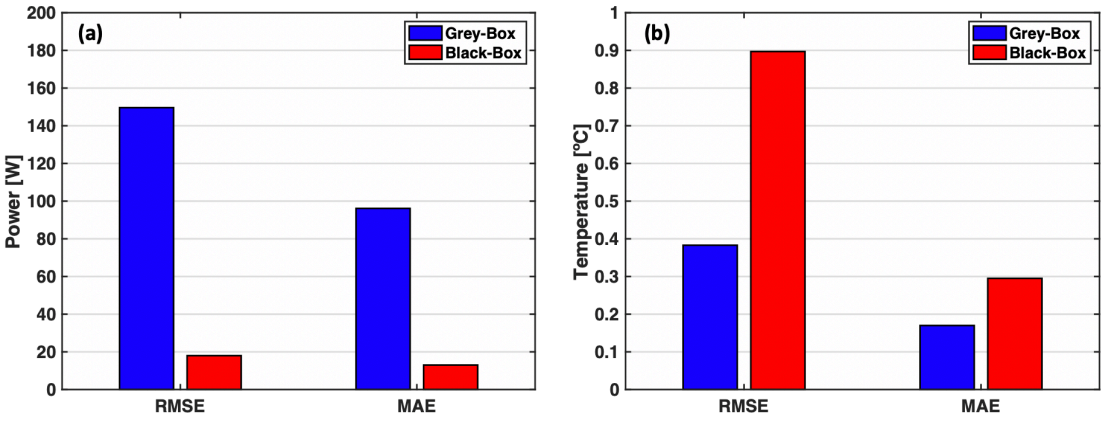

Figure 5.29: Comparison of grey-box and black-box modeling accuracy for (a) power and (b) temperature of water heater

the sources of error in the empirical data collected from the house described in Section 5.2. The black-box models outperformed the grey-box models in predicting the water heater power draw, but were worse than the grey-box models when predicting the water reservoir temperature, as shown in Figure 5.29. Despite the greater error of the black-box model when predicting the water reservoir temperature, the error of the black-box models was low. These results demonstrate that using black-box modeling methods are a viable alternative to grey-box methods. The black-box models may be used to model a variety of water heater and HVAC systems. Further analysis is needed to verify the performance of

the XGB algorithm on different appliances in a laboratory setting, but in theory, it may adapt to any electrical water heater or all-electric HVAC system. The inputs to the models are dynamic parameters such as flow rates and temperatures, which may be monitored in the same way for all appliances. The models infer any changes in appliance or house geometry that may influence system performance through training on data collected in the field.

CHAPTER 6. RESULTS

Optimization analyses were conducted for the months of August and January using both the black-box and grey-box models during the heating and cooling seasons. Time-of-use (TOU) pricing schemes for winter and summer were applied in the objective function to demonstrate the ability of the algorithm to respond to a price-driven demand response event. Initially, the grey-box models were optimized using a solver for disciplined linear problems as well as a novel genetic algorithm solver to demonstrate the advantages of the genetic algorithm. The genetic algorithm solver was then applied to the combinatorial optimization problem introduced by the use of black-box models. Multiobjective optimization was performed with the goal of maximizing energy and cost savings while minimizing any resulting thermal discomfort. Remaining within the minimum and maximum temperatures of the set point dead band for the water heater reservoir and indoor air was specified as the criterion for satisfying occupant thermal comfort requirements. The set points for the water heater and indoor air temperature for summer and winter periods are given in Table 6.1. The specified water heater set points are the set points during January and August of the actual water heater installed in the neighborhood. The HVAC set points are the set points of the HVAC system simulated in EES and used as a proxy for the actual HVAC system data in this simulation. The water heater and HVAC set points were constant for the duration of the respective simulation periods. The set point

Table 6.1: Water heating and space conditioning set points

	Water Heater		HVAC	
	Summer	**Winter**	**Summer**	**Winter**
Minimum set point temperature [ºC]	43.3	43.3	20	20
Set point temperature [ºC]	51.7	48.9	22	22
Maximum set point temperature [ºC]	60.0	60.0	23	23

temperature indicates the desired temperature of the system, while the minimum and maximum set point temperatures indicate the dead band in which the temperature may fluctuate.

6.1 Linear Optimization Solver

Initially, the MATLAB-based CVX solver (CVX Research Inc., 2012) was used to solve the constrained optimization problem formulated using the linear grey-box models and discussed in Chapter 4. This method is preferred due to its high computational speed (Jin et al., 2017), but it was found that the algorithm routinely resulted in increased cost and energy consumption while maintaining a minimal level of thermal comfort due to the optimization constraints and operational characteristics of the heat pump water heater. The optimization resulted in an increased electricity cost of 1% and increased energy consumption of 3% over a two-week simulation period. The optimized water heater reservoir was maintained at a mean temperature of 43.3ºC (equal to the minimum set point temperature) while the data measured in the field showed a mean water reservoir temperature of 49.4ºC. This was due to the solver attempting to maintain the minimum set point temperature of 43.3ºC, rather than the intermediate set point of 48.9ºC, as presented in Figure 6.1. The global minimum for the cost function is zero, but it is constrained by the upper and lower set points; therefore, the solver will converge on a solution along the boundary imposed by the minimum set point whenever possible, deviating from this boundary only when preheating to maintain the minimum set point for the following time step is required. The simulation was conducted using eight-hour time intervals with fifteen-minute control time steps. Maintaining the water reservoir temperature at the minimum set point temperature over one time interval resulted in a minimal initial temperature for the

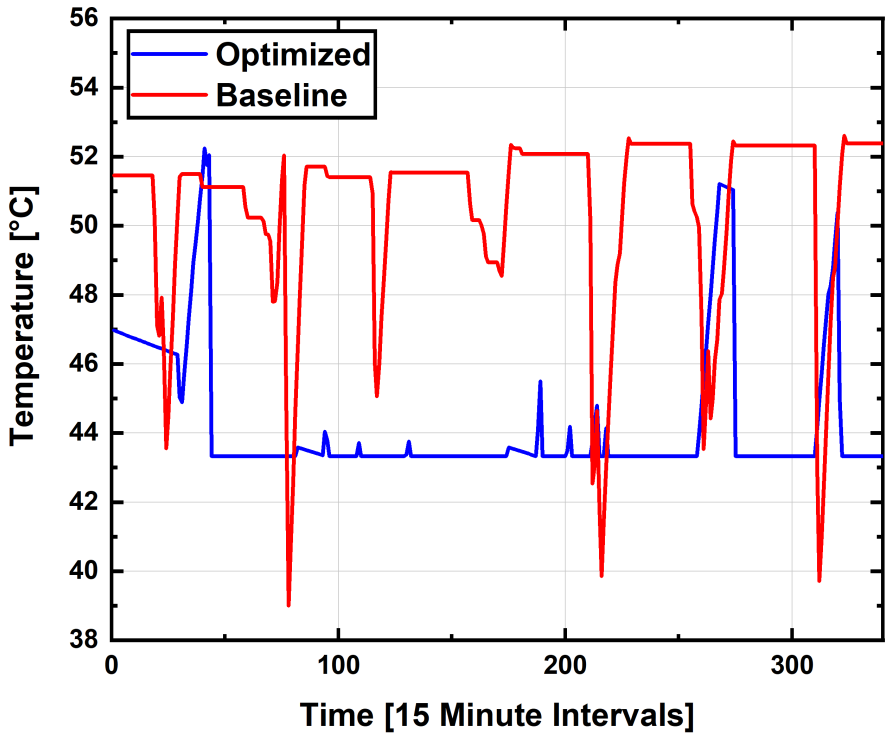

Figure 6.1: Baseline and optimized water reservoir temperatures using CVX solver

following interval. This low initial temperature led to situations in which the water heater was unable to keep the water temperature above the minimum set point threshold using the heat pump alone during time steps early in eight-hour prediction interval and immediately following large water draws. All water heater power draws greater than 500W in Figure 6.2 represent ON/OFF cycling events of the electric resistive elements. Activating the electric elements to meet the temperature requirement negated any energy savings resulting from the lower mean tank temperature. In addition, the optimization often resulted in infeasible solutions when the combination of electric resistive elements and heat pump were insufficient to meet the minimum temperature for the next time step. Infeasible solutions are those in which the optimal solution lies outside the feasible region as defined by the constraints of the function; in these instances, the solver does not return a solution.

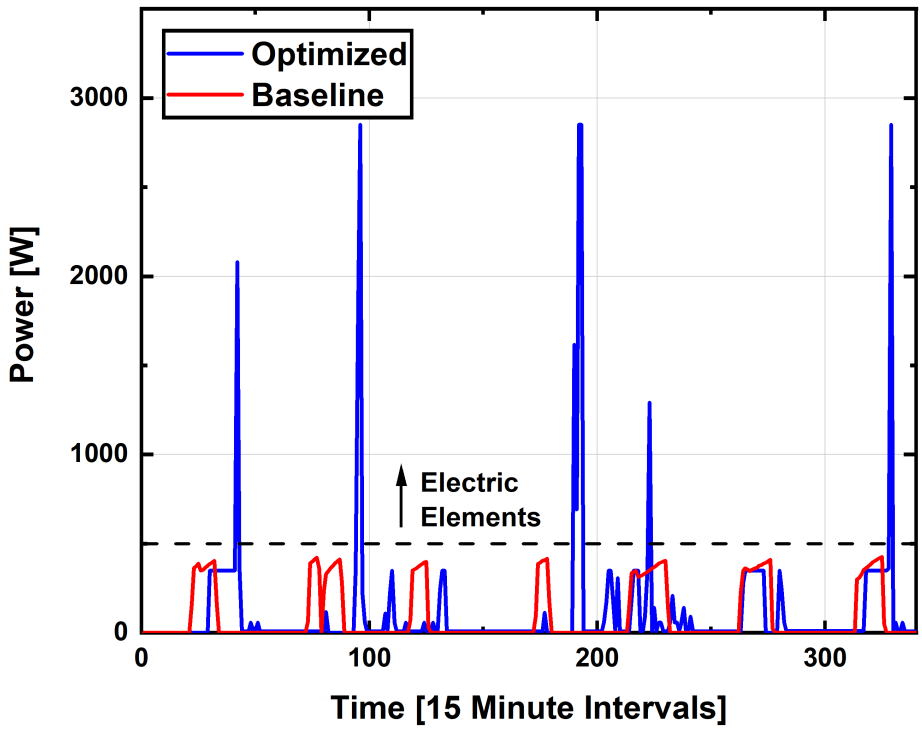

Figure 6.2: Water heater power draw for constrained linear optimization

Increasing the minimum set point to increase thermal comfort only led to more instances of infeasible solutions. It is expected that the issues encountered with the use of the constrained linear optimization solver would be a problem if the same approach was applied to any model of heat pump water heater with a similar or lower condenser heat duty, as the heat pump often does not provide the thermal energy necessary within a 15-minute time interval to recover the minimum set point temperature. To address these issues, a penalty method was applied in which temperatures deviating from the intermediate set point are penalized rather than constraining the temperature within a maximum and minimum bound (Yeniay, 2005; Konak et al., 2006). The results of applying this method are discussed in Section 6.2.

6.2 Grey-Box Models with Genetic Algorithm

The grey-box models were then used in an optimization algorithm that applies a genetic algorithm for the solver, as described in Chapter 4. This genetic algorithm was developed specifically to address the need for a combinatorial algorithm suitable for use with derivative-free black-box models but was also applied to the optimization of the grey-box models for results benchmarking. In addition to offering a method by which derivative-free optimization may be accomplished, the genetic algorithm allows penalties to be applied to the objective function rather than constraints, which solves the problems encountered when using disciplined solvers, as discussed in Chapter 4 and demonstrated in Section 6.1. The genetic algorithm was applied to the grey-box models in two scenarios:

- Scenario 1: Solar PV generation not included in the calculation for power drawn from the grid, P_{grid}.
- Scenario 2: Solar PV generation included in the calculation for power drawn from the grid, P_{grid}.

Scenario 1 allowed the appliance response to peak pricing periods to be isolated from the effect of solar PV generation to better understand the amount of load shifting or curtailment available from the controlled appliances. In Scenario 2, TOU pricing is combined with locally generated solar PV energy in the objective function. This scenario reflects the actual operating conditions in the neighborhood and demonstrates optimization with the goals of maximizing local solar PV energy consumption while minimizing thermal discomfort and electricity cost.

6.2.1 Multiobjective Weight Selection

To balance the prioritization of the various objectives included in the objective function, the optimal weight for each objective must be selected. One of the most challenging aspects of applying a weighted sum approach when formulating the objective function is choosing these optimal weight values (Konak et al., 2006). In the objective function, there are a total of eight weights, shown in Table 6.2, that are related to thermal comfort and may be varied, as well as the TOU price vector, which may be considered as a weight on the energy minimization objective. These weights may be varied to reflect the user's preference for various thermal comfort levels. For example, a user may prefer warmer indoor air temperatures, in which case, deviation above a maximum set point may be more desirable than deviation below a minimum setpoint, $\omega_5 > \omega_8$. The weights may also be varied to reflect preference for different optimization objectives. For example, exceedingly high price values may be set for the peak pricing period to reflect a desire to shift electrical load out of this period with little regard for thermal comfort, $c_e > \omega_n$.

In this study, positive and negative deviation from the set point values were given the same weight to reflect a desire to maintain temperatures within a set point dead band

Table 6.2: Optimization weights

Weight	Parameter
ω_1	Water temperature deviation > maximum set point
ω_2	Water temperature deviation > set point but < maximum set point
ω_3	Water temperature deviation < set point but > minimum set point
ω_4	Water temperature deviation < minimum set point
ω_5	Air temperature deviation > maximum set point
ω_6	Air temperature deviation > set point but < maximum set point
ω_7	Air temperature deviation < set point but > minimum set point
ω_8	Air temperature deviation < minimum set point
c_e	TOU pricing vector

with indifference toward warmer or colder preferences; i.e., for the water heater temperature, $\omega_1 = \omega_4$ and $\omega_2 = \omega_3$, and for the indoor air temperature, $\omega_5 = \omega_8$ and $\omega_6 = \omega_7$. The weights corresponding to temperatures that deviate from the set point without passing the minimum or maximum threshold $(\omega_2, \omega_3, \omega_6, \omega_7)$ were set to a constant value of 0.25, and the price vector was considered to be static for each season. Two weights that could be varied to find an optimal set remained; these weights are referred to as ω_{WH} and ω_{HVAC}, where $\omega_{WH} = \omega_1 = \omega_4$ and $\omega_{HVAC} = \omega_5 = \omega_8$. To find an optimal set of weights, a parametric study was conducted to evaluate the set of Pareto optimal solutions generated by varying the weights over a range of ω_{WH} and ω_{HVAC}. The weights were initially varied iteratively over a wide range of values. It was found that weights > 4 for the water heater and > 3 for the HVAC system during the month of August resulted in maximum priority being placed on thermal comfort with no cost or energy savings. ω_{HVAC} was varied in increments of 0.5 from 0.5 to 4, and ω_{WH} was varied in increments of 0.5 from 0.5 to 3. The results were evaluated based on the resulting daily cost and a thermal comfort value. The thermal comfort value was determined by summing the temperature difference from the setpoint for both the water heater reservoir and indoor air for all timesteps over the course of a day, using Equation (6.1).

$$\text{Daily Set Point Deviation} = \sum_{t}^{T} \left(\left| T_{WH}(t) - T_{WH}^{SP}(t) \right| + \left| T_{air}^{in}(t) - T_{air}^{SP}(t) \right| \right) \tag{6.1}$$

Figure 6.3 shows the results of this study for August; weights of $\omega_{WH} = 2$ and $\omega_{HVAC} = 3$ were chosen for summer months by selecting the point at which the lowest daily cost for the lowest daily set point deviation was achieved.

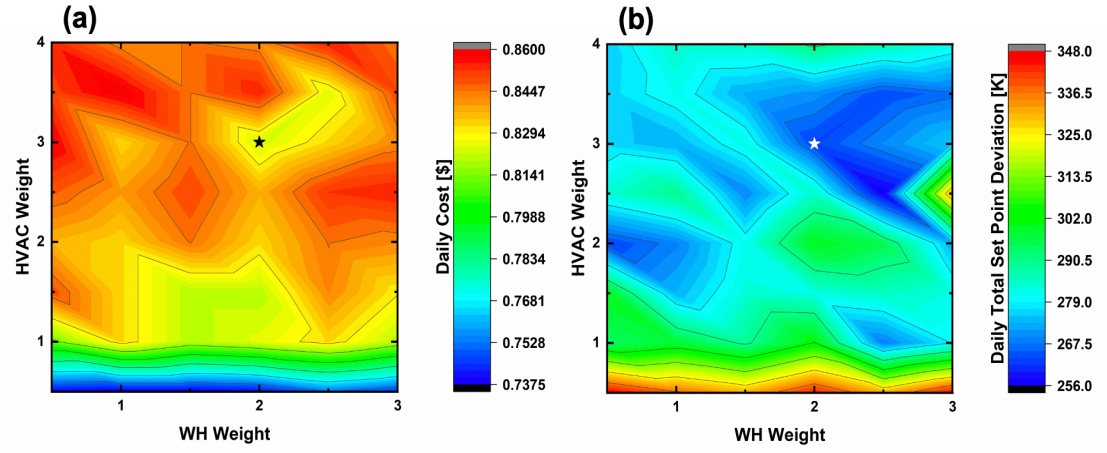

Figure 6.3: Weight optimization cost and thermal comfort results (summer)

The same analysis was performed for January, with ω_{WH} varied from 1 to 3 and ω_{HVAC} varied from 1 to 10. Initially, the range of weights investigated for January was the same as that of August, but the increase in HVAC energy consumption and difference in the price vector resulted in a comparatively homogeneous plot region with no readily apparent optimal choice of weights. The range of weights for the HVAC system was extended through ten to investigate other regions in which a better weight combination could be found. The resulting contour plots are presented in Figure 6.4. Weights of $\omega_{WH} = 1$ and $\omega_{HVAC} = 4$ were chosen for January.

For this optimization scenario, the weight combination for winter and summer that balances the goals of cost minimization and thermal comfort was selected. However, other weights may be selected depending on the desired outcome. For example, a user may decide to prioritize thermal comfort over all else, at the expense of energy and cost savings. Figure 6.3 and Figure 6.4 demonstrate the effect on daily cost and thermal comfort caused by varying the weights $\omega_{WH} = \omega_1 = \omega_4$ and $\omega_{HVAC} = \omega_5 = \omega_8$ for the HVAC system and water heater. Inspecting these contour plots gives insight into the behavior of the system in

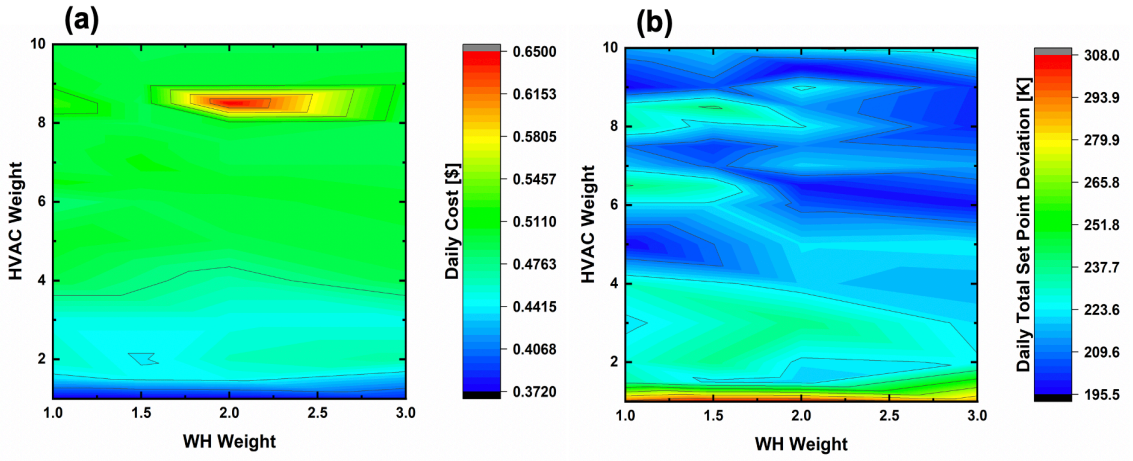

Figure 6.4: Weight optimization cost and thermal comfort results (winter)

response to a given weight combination. For example, choosing the maximum weight in Figure 6.3 for both the water heater and HVAC system results in the lowest level of total deviation from the set point, indicating that these weights result in the highest thermal comfort levels, but also the highest cost. More sophisticated methods of incorporating the weighting problem into the algorithm are described in the literature (Konak et al., 2006), but were not incorporated in this version of the algorithm due to computational time constraints.

6.2.2 Scenario 1: TOU Pricing Without Solar PV

In the first scenario, solar PV was excluded from the optimization. Considering the effects of optimization with only the effect of TOU pricing included in the objective function allowed for the analysis of the load shifting capability of the appliances without the influence of solar PV generation. The results of the simulation are presented in Table 6.3 and Figure 6.5. In this table, total energy is the total amount of energy consumed by the house for the month. Peak energy is the total amount of energy consumed by the house during peak pricing hours over the month. Total cost is the electricity cost for the month.

Table 6.3: August grey-box optimization results (no PV generation)

	Baseline	Optimized	% Reduction, *r*
Total energy [kWh]	681.75	649.04 ± 1.42	4.6 < *r* < 5.0
Peak energy [kWh]	212.76	184.91 ± 1.42	12.4 < *r* < 13.8
Total cost [$]	84.54	77.85 ± 0.17	7.7 < *r* < 8.1
HVAC energy [kWh]	139.81	101.44 ± 0.03	27.4 < *r* < 27.5
Peak HVAC energy [kWh]	54.46	33.09 ± 0.01	39.2 < *r* < 39.3
WH energy [kWh]	40.59	49.93 ± 1.27	-26.1 < *r* < -19.9
Peak WH energy [kWh]	7.06	1.63 ± 0.29	72.8 < *r* < 81.0

HVAC energy and WH energy are the total amounts of energy consumed by the HVAC system and water heater during the month. Peak HVAC energy and peak WH energy are the total amounts of energy consumed by the HVAC system and water heater during peak pricing hours over the month. Uncertainty was calculated for the optimized results and compared with the baseline, with the baseline assumed as a datum reference. Sample

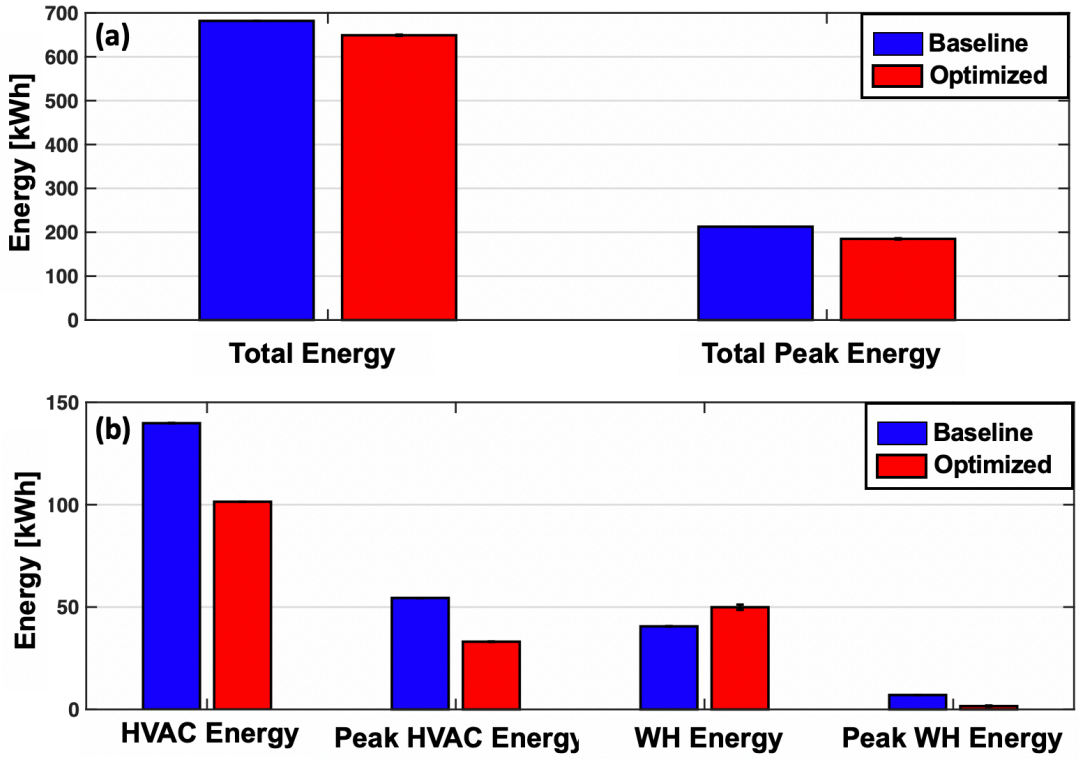

Figure 6.5: August grey-box optimization results with no PV energy generation for (a) whole-house energy, (b) appliance energy

uncertainty calculations may be found in Appendix B. Results show that total energy consumption was reduced by up to 5% in the month of August. Total peak energy for the month of August decreased by 12.4 to 13.8% as a result of controllable electrical load being shifted out of the peak pricing period and the combination of peak shifting and total energy reduction resulted in a cost savings of \$6.69 $\pm$ \$0.17 for the customer over the month.

The total energy reduction can be attributed to reduced operation of appliances, resulting in the temperatures presented in Table 6.4 and Figure 6.6. While the total energy consumption of the house was reduced, the total energy consumption of the water heater increased by between 19.9 and 26.1%. The increase in total energy consumption of the water heater can be attributed to the maintenance of thermal comfort while shifting load out of the peak period. The water heater temperature is more volatile than the indoor air temperature due to the rapid temperature changes caused by water draws. Therefore, the tank often preheats prior to a peak demand event to maintain the set point temperature during the peak period without operating the heat pump or electric resistive elements. Preheating the tank uses more energy outside of the peak period to heat the tank to a temperature greater than what would normally be seen approaching the peak pricing period, so that the water heater can avoid consuming energy during the peak period. In addition, the electric resistive elements are occasionally activated immediately after the peak period

Table 6.4: August water reservoir and indoor air temperatures (no PV generation)

	Water Reservoir Temperature [ºC]		Indoor Air Temperature [ºC]	
	Baseline	**Optimized**	**Baseline**	**Optimized**
Mean	51.48	50.42 $\pm$ 9.12	21.82	22.39 $\pm$ 3.23
Maximum	53.08	53.21 $\pm$ 0.17	22.98	23.24 $\pm$ 0.06
Minimum	45.55	45.13 $\pm$ 0.17	19.77	20.72 $\pm$ 0.06

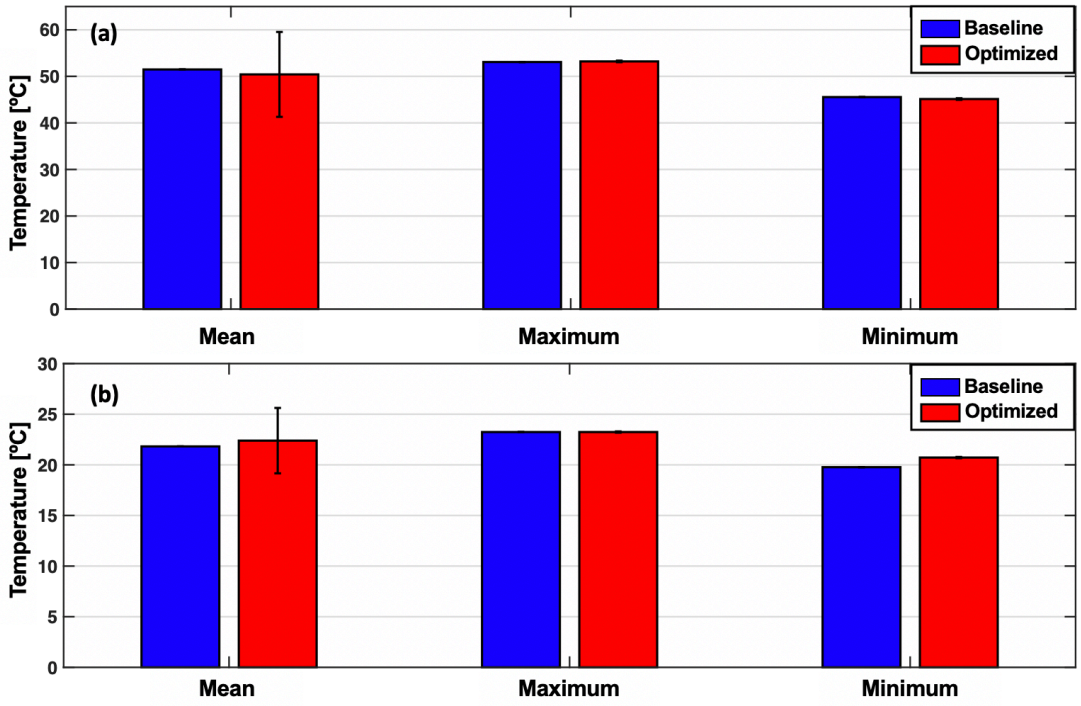

Figure 6.6: August (a) water reservoir and (b) indoor air temperatures with no PV generation

to recover the water temperature. The optimized water reservoir temperature never exceeded the maximum set point temperature or fell below the minimum set point temperature. The high uncertainty associated with the optimized mean water heater temperature makes it a poor figure for comparison, but the maximum and minimum water reservoir temperatures are approximately equal to those observed during baseline operation. The high mean temperature uncertainty is a result of model error propagation over time. An example of this calculation may be found in Appendix B. The baseline mean indoor air temperature was within the range of the optimized indoor air temperature. Both the minimum and maximum indoor air temperatures exceeded those observed in the baseline case. The increase in indoor temperature and reduction in peak HVAC energy consumption demonstrates the sacrifice of thermal comfort during cooling season for peak load shifting. Optimization outcomes for the month of January are reported in Table 6.5

Table 6.5: January grey-box optimization results (no PV generation)

	Baseline	Optimized	% Reduction, *r*
Total energy [kWh]	821.25	797.16 ± 0.26	2.9 < *r* < 3.0
Peak energy [kWh]	146.82	104.71 ± 0.59	28.3 < *r* < 29.1
Total cost [$]	66.42	63.73 ± 0.11	3.8 < *r* < 4.2
HVAC energy [kWh]	306.57	295.47 ± 0.03	3.6
Peak HVAC energy [kWh]	57.95	23.66 ± 0.01	59.2
WH energy [kWh]	53.08	50.24 ± 1.26	3.0 > *r* > 7.7
Peak WH energy [kWh]	10.83	3.29 ± 0.42	65.7> *r* > 73.5

and Figure 6.7. The resulting temperatures are reported in Table 6.6 and Figure 6.8. In January, total energy and peak energy decreased by 2.9 – 3.0% and 28.3 – 29.1%, respectively; the peak and total energy consumption of both the HVAC and water heater decreased as a result of optimization during January. Total cost decreased by 3.8 - 4.2%. The optimized minimum and maximum water reservoir temperatures are higher than the baseline temperatures, which suggests that greater energy savings may be possible by

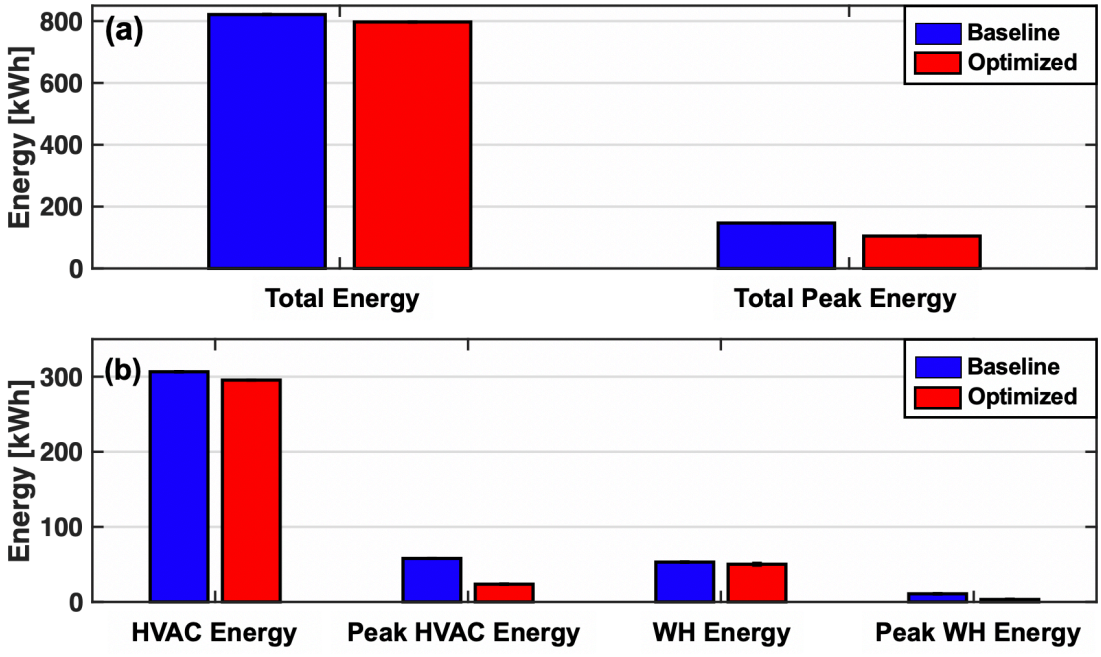

Figure 6.7: January grey-box optimization results with no PV generation for (a) whole-house energy, (b) appliance energy

Table 6.6: January water reservoir and indoor air temperatures (no PV generation)

	Water Reservoir Temperature [ºC]		Indoor Air Temperature [ºC]	
	Baseline	**Optimized**	**Baseline**	**Optimized**
Mean	45.25	46.34 ± 9.12	21.64	21.43 ± 3.23
Maximum	47.71	51.14 ± 0.17	24.48	24.89 ± 0.06
Minimum	36.54	40.31 ± 0.17	19.50	18.30 ± 0.06

varying ω_{WH} in such a way that the water heater reservoir is maintained at a temperature equal to or lower than the baseline reservoir temperatures. The mean indoor air temperature followed the inverse of the trend of the indoor air temperature observed in August, which accounts for the difference in cooling and heating season operation. The mean baseline indoor air temperature in January was within the range of the optimized mean temperature, the maximum temperature was greater than the maximum baseline temperature, and the minimum optimized temperature was lower than the minimum baseline temperature. This

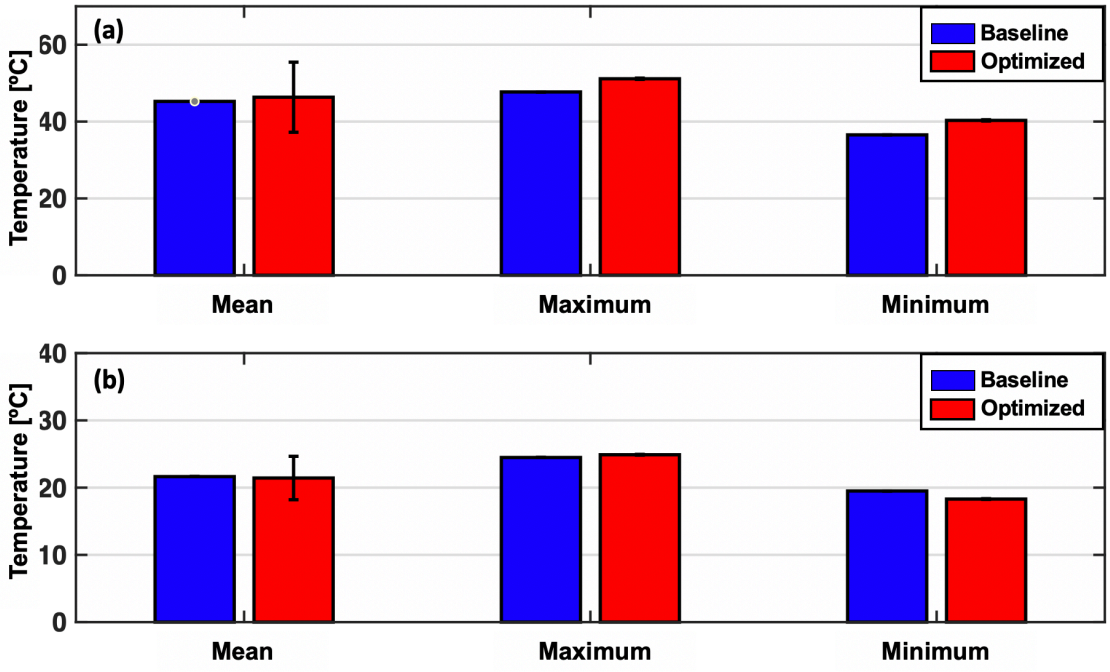

Figure 6.8: January (a) water reservoir and (b) indoor air temperatures with no PV generation

indicates that the optimization algorithm allowed the indoor air temperature to fluctuate within a wider temperature band to shift load outside of peak pricing periods and reduce overall energy consumption while maintaining thermal comfort within the specified set point dead band.

6.2.3 Scenario 2: TOU pricing and solar PV

Adding solar PV generation to the optimization further enhanced energy and cost savings by allowing load shifted out of the peak pricing period to be shifted into the peak solar PV generation period. In this scenario, total and peak energy represent the net energy consumed by the house and include energy generated by solar PV. Results of the optimization for the month of August are presented in Table 6.7 and Figure 6.9. Optimization resulted in increases in the total and peak energy consumption of the water heater for the month of August, even though total energy consumed by the house was reduced. These increases are due to the control actions of the water heater in response to the high levels of solar PV energy generation available during August. The water heater consumed as much energy as necessary to maintain its set point during peak solar PV generation hours, due to there being more PV energy supplied than can be used by the house.

Table 6.7: August grey-box optimization results (with PV generation)

	Baseline	Optimized	% Reduction, *r*
Total energy [kWh]	359.15	316.11 ± 3.15	11.1 < *r* < 12.9
Peak energy [kWh]	39.49	26.69 ± 0.62	30.8 < *r* < 34.0
Total cost [$]	33.65	28.41 ± 0.35	14.5 < *r* < 16.6
HVAC energy [kWh]	145.36	116.02 ± 0.03	20.1 < *r* < 20.2
Peak HVAC energy [kWh]	56.57	50.04 ± 0.01	11.5 < *r* < 11.6
WH energy [kWh]	41.60	73.41 ± 1.29	-79.6 < *r* < -73.3
Peak WH energy [kWh]	8.35	9.00 ± 0.51	-13.9 < *r* < -1.7

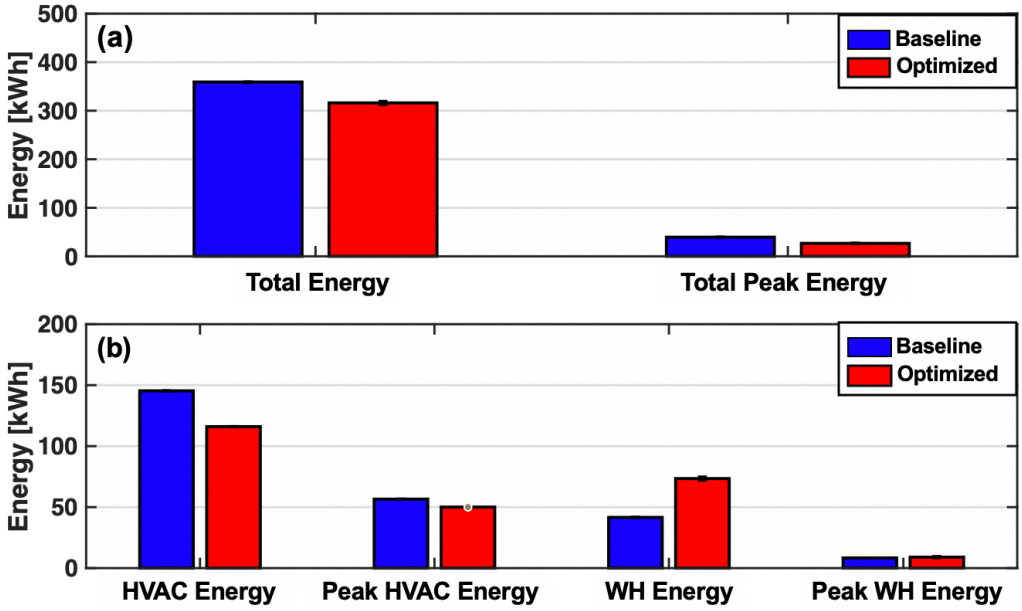

Figure 6.9: August grey-box optimization results with PV generation for (a) whole-house energy, (b) appliance energy

The water reservoir and indoor air temperatures for the month of August are presented in Table 6.8 and Figure 6.10. The minimum water reservoir temperature for this optimization scenario is higher than the mean minimum temperature of the optimized tank without solar PV generation, demonstrating that the tank is using more energy to maintain a higher internal temperature. The maximum and minimum indoor air temperatures fall outside the set point dead band range in both the baseline and optimized cases; increasing the value of ω_{HVAC} would reduce or eliminate deviation from the set point dead band result in greater energy consumption.

Table 6.8: August water reservoir and indoor air temperatures (with PV generation)

	Water Reservoir Temperature [ºC]		Indoor Air Temperature [ºC]	
	Baseline	**Optimized**	**Baseline**	**Optimized**
Mean	51.53	51.21 ± 9.12	21.81	22.21 ± 3.23
Maximum	53.09	53.19 ± 0.17	22.98	22.99 ± 0.06
Minimum	45.72	46.31 ± 0.17	19.77	20.55 ± 0.06

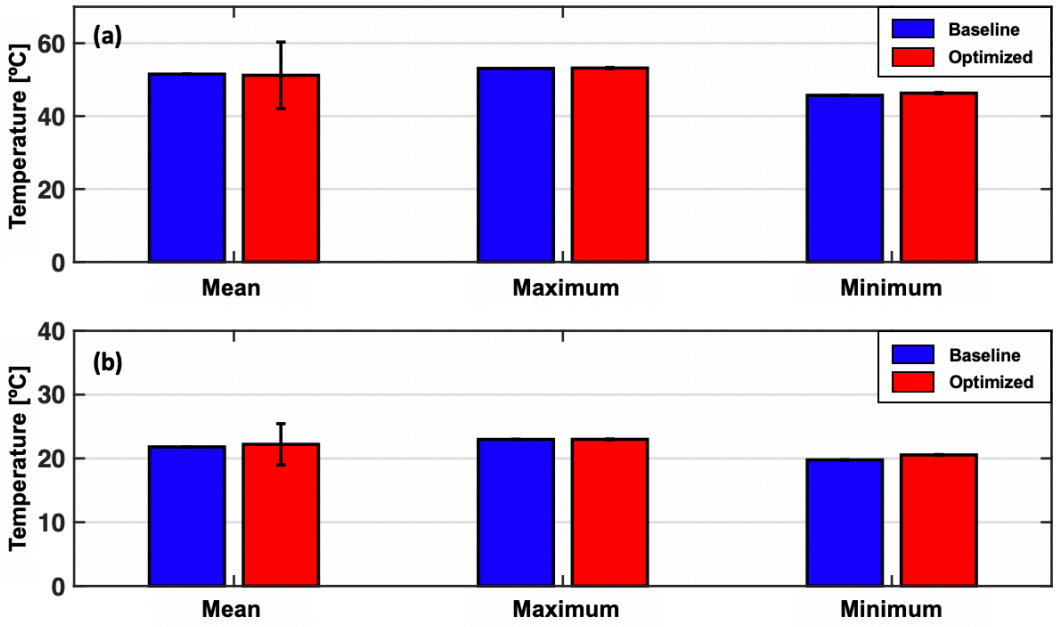

Figure 6.10: August (a) water reservoir and (b) indoor air temperature with PV generation

Results of the optimization with solar PV generation for the month of January are presented in Table 6.9 and Figure 6.11 and the resulting temperatures are presented in Figure 6.12 and Table 6.10. The percent reduction for total energy and peak energy as compared to the baseline was greater for the case with PV generation than without, but the percent cost reduction was less. This is due to the reduction in baseline energy consumption.

Table 6.9: January grey-box optimization results (with PV generation)

	Baseline	Optimized	% Reduction, *r*
Total energy [kWh]	621.05	597.57 ± 11.38	$1.9 < r < 5.6$
Peak energy [kWh]	102.26	61.51 ± 1.44	$38.4 < r < 41.3$
Total cost [$]	55.28	52.99 ± 0.91	$2.5 < r < 5.8$
HVAC energy [kWh]	306.57	299.55 ± 0.003	2.3
Peak HVAC energy [kWh]	57.95	25.31 ± 0.001	56.3
WH energy [kWh]	53.78	75.76 ± 0.47	$-41.7 < r < -40.0$
Peak WH energy [kWh]	10.83	6.55 ± 2.27	$18.6 < r < 60.5$

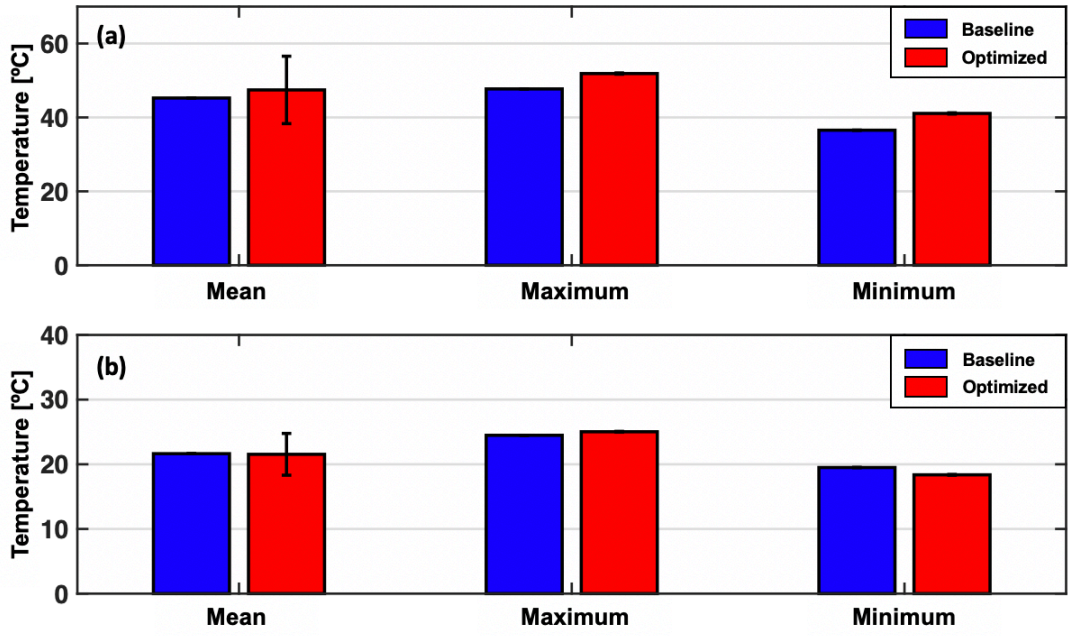

Figure 6.11: January (a) water reservoir and (b) indoor air temperatures with PV generation

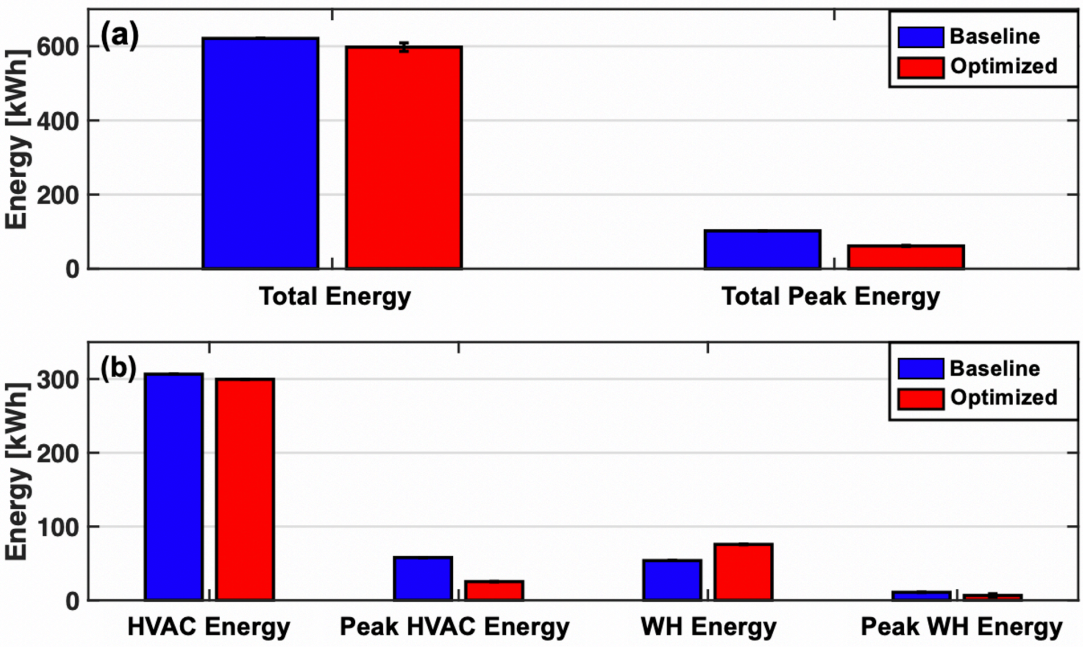

Figure 6.12: January grey-box optimization with PV generation for (a) whole-house energy, (b) appliance energy

Table 6.10: January water reservoir and indoor air temperatures

	Water Reservoir Temperature [ºC]		Indoor Air Temperature [ºC]	
	Baseline	**Optimized**	**Baseline**	**Optimized**
Mean	45.25	47.44 ± 9.12	21.64	21.54 ± 3.23
Maximum	47.71	51.87 ± 0.17	24.48	25.03 ± 0.06
Minimum	36.54	41.08 ± 0.17	19.50	18.37 ± 0.06

Both the water reservoir and indoor air temperature during January were maintained at higher temperatures as compared to the optimization scenario for the month of January with no PV generation, due to the excess PV energy generation. This trend was also seen in the month of August, where excess PV generation allowed more energy to be used for water heating and indoor air cooling.

6.3 Black-Box Models with Genetic Algorithm

The genetic algorithm was then applied to the period of the day in which the peak scheduling occurs for both August and January using the black-box models, to demonstrate the agreement of the outcomes between optimization using the two modeling methods.

The genetic algorithm was then applied to the period of the day in which the peak scheduling occurs for both August and January using the black-box models, to demonstrate the agreement of the outcomes between optimization using the two modeling methods. TOU peak pricing periods and solar PV generation are included in this optimization. Table 6.11 and Figure 6.13 presents the results of the optimization using black-box models compared with the optimization results using grey-box models for the month of August. After considering modeling uncertainty, the reduction in total and peak energy for the black-box and grey-box models are approximately the same. In both approaches, the water heater and HVAC system used more energy as compared to the baseline case, with the

Table 6.11: August optimization results (single day)

		Baseline	Optimized	% Reduction, r
Black-Box Optimization Results	**Total energy [kWh]**	0.93	0.67 ± 0.01	26.9 < r < 29.0
	Peak energy [kWh]	0.93	0.67 ± 0.01	26.9 < r < 29.0
	Total cost [$]	0.21	0.15 ± 0.02	19.0 < r < 38.1
	HVAC energy [kWh]	1.96	2.37 ± 0.002	-21.0 < r < -20.8
	Peak HVAC energy [kWh]	1.19	1.60 ± 0.001	-34.5 < r < -34.4
	WH energy [kWh]	0.62	2.01 ± 0.02	-227 < r < -221
	Peak WH energy [kWh]	0.00	1.24 ± 0.01	-
Grey-Box Optimization Results	**Total energy [kWh]**	0.93	0.69 ± 0.06	19.4 < r < 32.3
	Peak energy [kWh]	0.93	0.69 ± 0.08	17.2 < r < 34.4
	Total cost [$]	0.21	0.16 ± 0.03	9.5 < r < 38.1
	HVAC energy [kWh]	1.96	2.53 ± 0.003	-29.2 < r < -28.9
	Peak HVAC energy [kWh]	1.19	1.95 ± 0.002	-64.0 < r < -63.7
	WH energy [kWh]	0.62	1.23 ± 0.14	-121 < r < -75.8
	Peak WH energy [kWh]	0.00	0.10 ± 0.10	-

black-box optimization approach using more energy for both appliances than the grey-box approach. The resulting temperatures are presented in Table 6.12 and Figure 6.14, where it may be noted that the greater energy consumption of the black-box approach has resulted

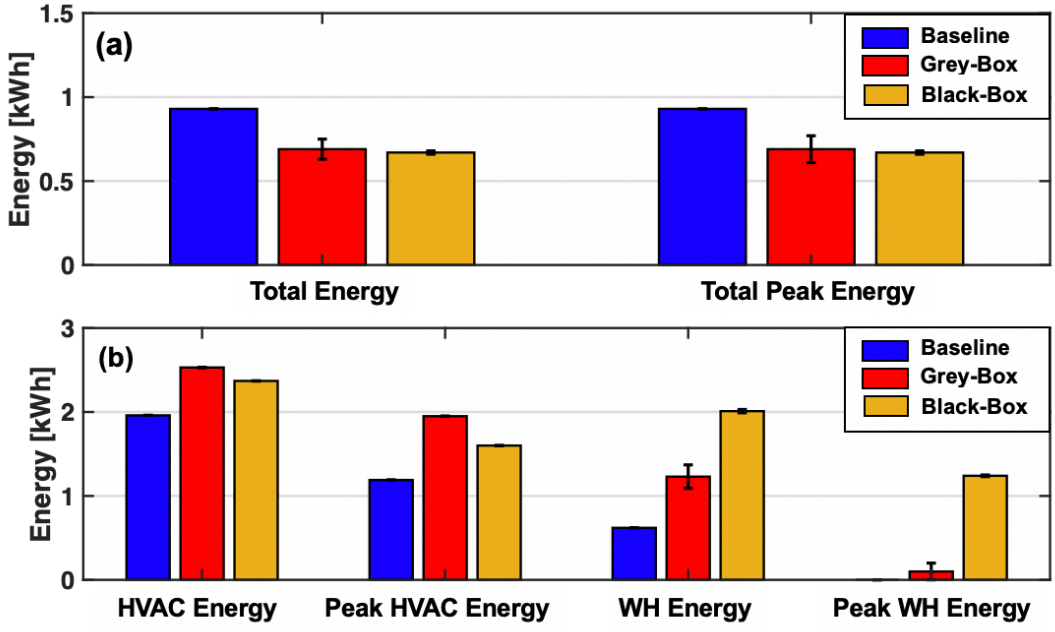

Figure 6.13: August grey-box optimization results (a) whole-house energy, (b) appliance energy

Table 6.12: August water reservoir temperatures comparison (single day)

Water Reservoir Temperature [ºC]			
	Baseline	**Black-Box Optimization**	**Grey-Box Optimization**
Mean	51.65	52.64 ± 2.19	51.40 ± 0.96
Minimum	48.37	49.94 ± 0.39	50.21 ± 0.17
Maximum	52.48	54.43 ± 0.39	52.25 ± 0.17
Indoor Air Temperature [ºC]			
	Baseline	**Black-Box Optimization**	**Grey-Box Optimization**
Mean	21.72	21.95 ± 0.26	21.60 ± 0.34
Minimum	19.94	20.42 ± 0.05	20.05 ± 0.06
Maximum	23.00	22.99 ± 0.05	23.07 ± 0.06

in higher water reservoir temperatures and a lower minimum indoor air temperature. The disparity in appliance energy consumption values for the two controllable appliances may be explained by once more considering the effect of excess PV generation. Because the objective function is designed to minimize the total energy consumption of the house and net energy metering is not considered, any excess PV generation may be consumed by the

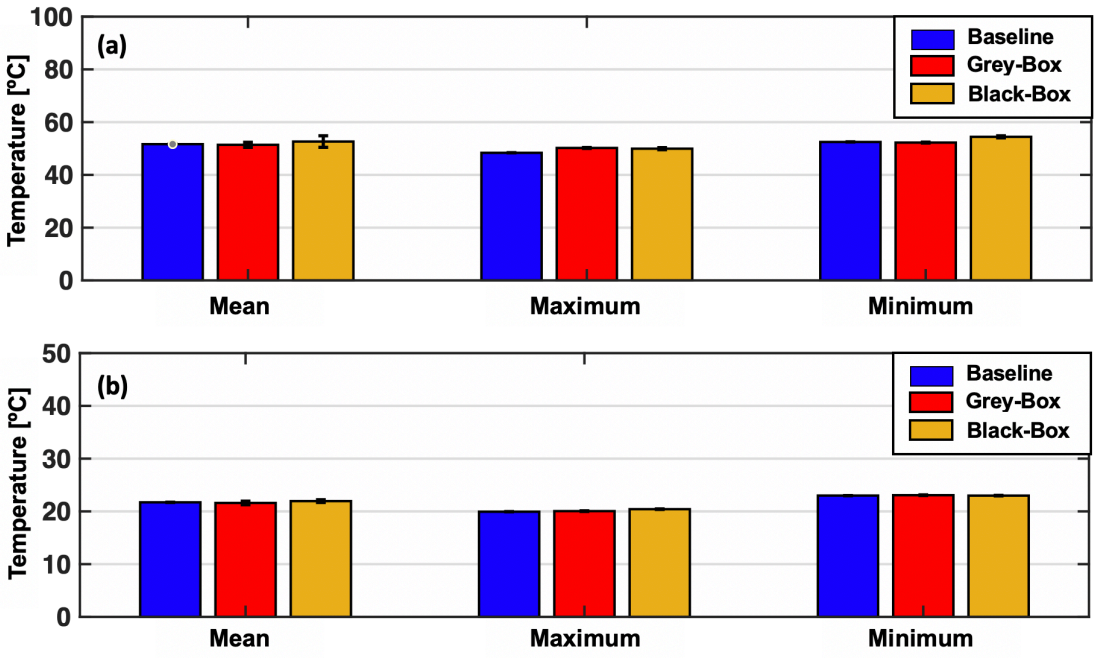

Figure 6.14: August (a) water reservoir and (b) indoor air temperatures

individual appliances with no cost penalty. The only associated penalty for appliance operation in this case, where P_{grid} is equal to zero, is the penalty for violating the thermal comfort dead band.

Table 6.13 and Figure 6.15 presents the results of the black-box and grey-box optimization methods for a day in January. In this case, total energy and cost savings were lower due to the lower peak energy prices and reduced solar PV generation. Considering the prediction uncertainty, the peak energy reduction of the two models is approximately equal, but the energy consumption of the grey-box model optimization is slightly greater than that of the black-box model optimization. The resulting temperatures presented in Figure 6.16 and Table 6.14 show that the black-box method produces temperatures more similar to the baseline temperatures than those that the grey-box method produces, which is expected considering the greater energy consumption associated with the water heater in the grey-box method.

Table 6.13: January optimization results (single day)

		Baseline	Optimized	% Reduction, *r*
Black-Box Optimization Results	**Total energy [kWh]**	5.06	4.57 ± 0.2	5.7 < *r* < 13.6
	Peak energy [kWh]	3.99	2.78 ± 0.12	27.3 < *r* < 33.3
	Total cost [$]	0.47	0.39 ± 0.01	14.9 < *r* < 19.2
	HVAC energy [kWh]	1.44	1.46 ± 0.01	-2.1 < *r* < -0.7
	Peak HVAC energy [kWh]	1.44	0 ± 0.01	99.3 < *r* < 100
	WH energy [kWh]	0.98	0.60 ± 0.02	36.7 < *r* < 40.8
	Peak WH energy [kWh]	0.51	0.15 ± 0.01	68.6 < *r* < 72.5
Grey-Box Optimization Results	**Total energy [kWh]**	5.06	5.12 ± 0.45	-10.1 < *r* < 7.7
	Peak energy [kWh]	3.99	2.67 ± 0.27	26.3 < *r* < 39.8
	Total cost [$]	0.47	0.43 ± 0.06	-4.3 < *r* < 21.3
	HVAC energy [kWh]	1.44	1.44 ± 0.003	-0.2 < *r* < 0.2
	Peak HVAC energy [kWh]	1.44	0 ± 0.002	99.9 < *r* < 100
	WH energy [kWh]	0.98	1.22 ± 0.14	-38.8 < *r* < -10.2
	Peak WH energy [kWh]	0.51	0 ± 0.10	80.4 < *r* < 100

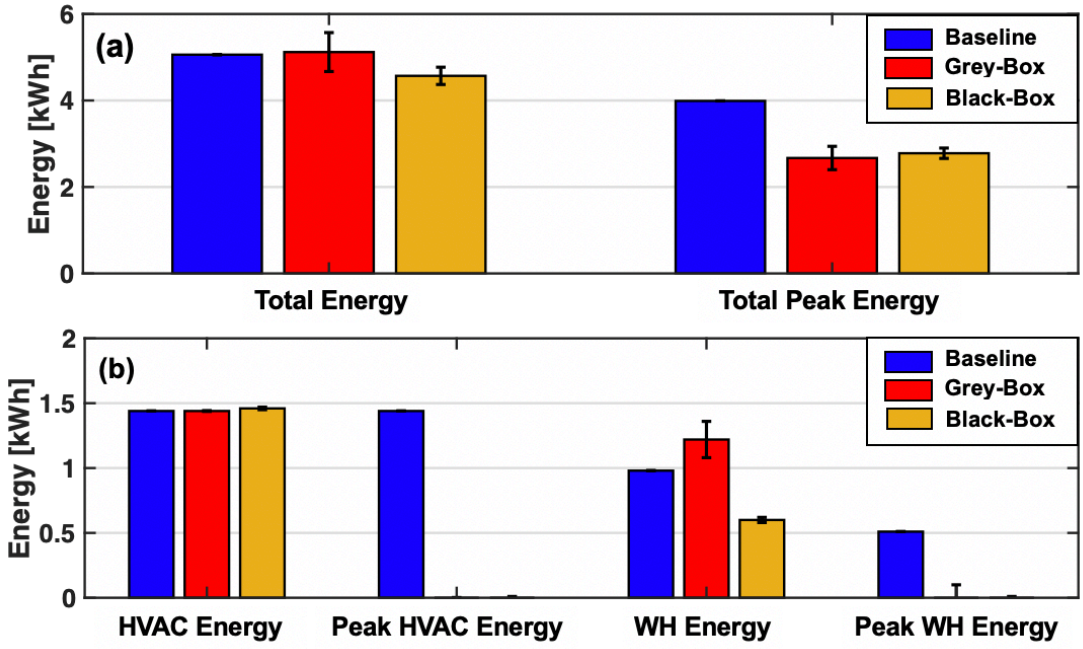

Figure 6.15: January optimization results (a) whole-house energy, (b) appliance energy

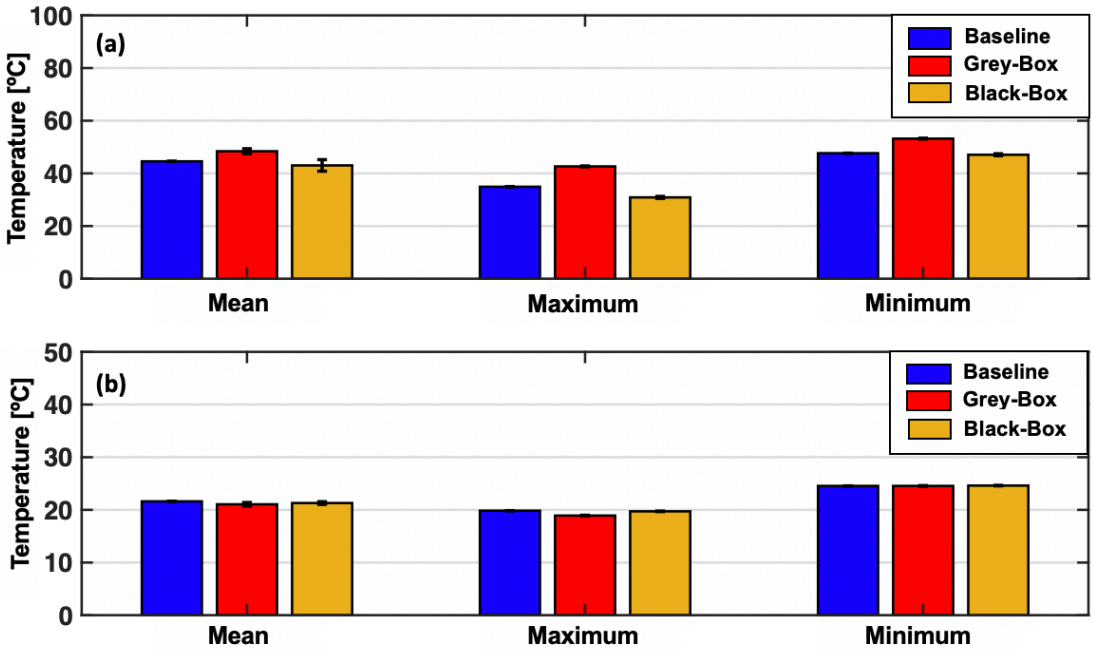

Figure 6.16: January (a) water reservoir and (b) indoor air temperatures

Table 6.14: January water reservoir temperatures comparison (single day)

Water Reservoir Temperature [ºC]			
	Baseline	**Black-Box Optimization**	**Grey-Box Optimization**
Mean	44.57	43.02 ± 2.19	48.40 ± 0.96
Minimum	34.92	30.88 ± 0.39	42.65 ± 0.17
Maximum	47.67	47.07 ± 0.39	53.20 ± 0.17
Indoor Air Temperature [ºC]			
	Baseline	**Black-Box Optimization**	**Grey-Box Optimization**
Mean	21.62	21.30 ± 0.26	21.06 ± 0.34
Minimum	19.85	19.72 ± 0.05	18.92 ± 0.06
Maximum	24.55	24.61 ± 0.05	24.55 ± 0.06

6.4 Results Summary

The disciplined convex solver CVX was applied to the optimization problem formulated using the grey-box modeling method to illustrate issues associated with disciplined programming that are unique to the optimization of the energy consumption of a heat pump water heater. The optimization problem was then solved using a genetic algorithm with the same grey-box models to demonstrate the effectiveness of the genetic algorithm compared to the disciplined linear solver and to provide benchmarking results for the analysis of the genetic algorithm using black-box models. In addition to allowing system models to be generated with limited user input, the black-box models of the water heater and HVAC system were shown in Chapter 5 to outperform the grey-box models due to their ability to infer nonlinear and transient system behavior that is not captured by simplified linear models. The genetic algorithm optimization using black-box modeling methods was demonstrated to provide cost and energy savings similar to that found using the grey-box modeling method, but with less inputs required from the user.

The maximum relative amount of total house energy savings achieved using this optimization was 12.9%, observed in August with the inclusion of PV energy. The mean relative total energy savings was 5.88%. This result is similar to mean energy savings of 4.56% reported by Jin et al. (2017), who used grey-box models similar to those developed in the present study. The maximum relative amount of peak energy shifted was 41.3%, observed in January with the inclusion of PV energy. The mean amount of peak energy shifted was 28.5%. In all cases, the HVAC energy was reduced, while the water heater energy stayed the same or increased. This may be explained by the greater ability of the water heater to preheat prior to demand response events. With a wider temperature dead band, the water heater has much greater flexibility than the few degrees of dead band fluctuation allowed by the HVAC system. This led to the water heater frequently preheating prior to the peak demand period or preheating during peak solar PV hours. This increased energy consumption of the water heater but resulted in overall energy and cost savings. The algorithm prioritized minimizing HVAC system operation rather than preheating or precooling the space, which resulted in lower HVAC energy consumption. The results reported may vary depending on occupant behavior and house characteristics. For example, occupants may not prioritize thermal comfort and allow the HVAC system a wider dead band in which it may operate, resulting in more preheating and precooling events. The houses in this study are very energy efficient. Older homes with more infiltration, less insulation, and more inefficient equipment, could be expected to experience greater reductions in absolute energy consumption using optimization than those reported in this study.

The maximum cost savings reported in this study were $6.86 for the month of August and $3.20 for the month of January. Such low cost savings may not be balanced by the sacrifice in thermal comfort made by customers. The cost reduction in this study was the result of using a time-of-use pricing schedule, but other pricing methods may be used in practice. The pricing schedule may be altered to incentivize customers to shift load out of peak demand periods. This may be accomplished using a day-ahead real-time pricing schedule that is transmitted to the house HEMS using the utility AMI meter.

The choices of weights for the HVAC and water heater also had an impact on the optimization results. In this study, a weight combination was selected by the author from the set of Pareto optimal solutions resulting from various weight combinations. In practice, the weight combination would be selected based on consumer preference for various objectives, including cost savings and thermal comfort, after evaluating the Pareto front. In other words, the Pareto front would be evaluated and a ranking of the resulting solutions would then be performed using customer preference. This customer preference ranking could be performed using a method such as the Simple Multi-Attribute Rating Technique Exploiting Ranks (SMARTER), as described in Jin et al. (2017).

Computational time for the genetic algorithm was rather long. For the genetic algorithm using grey-box models, optimization of an 8-hour control window required three hours of computational time on a MacBook Pro with a 2.4 GHz Intel i9 processor and 32 GB 2400 MHz DDR4 memory. The genetic algorithm optimization with black-box models required approximately six hours of computational time. Computational time results will vary depending on machine specifications such as processor speed and memory capacity. Increasing the speed of the algorithm may be the topic of future research.

CHAPTER 7. CONCLUSION

An optimization algorithm with the goals of reducing electricity cost for residential consumers, maximizing the local dispatch of on-site solar PV energy, and shifting load out of peak demand periods while maintaining customer comfort was developed. The algorithm was applied to a 2876 ft^2 (267.2 m^2) house constructed using energy-conserving building practices and equipped with energy efficient appliances. The house is estimated to consume 40-50% of the energy consumed by a typical newly constructed house in the United States and is located in a neighborhood consisting of sixty-two similarly constructed houses. An on-site community-level solar PV generation facility supplies the houses with renewable energy. For the purpose of this study, the PV generation facility is assumed to serve only this house, and its production was scaled by 1/62 to reflect this. Grey-box and black-box models were developed to predict the power draw of a heat pump water heater and HVAC system given a set of environmental conditions relevant to system operation. A second set of grey-box and black-box models were then developed to predict the temperature change of the water reservoir and indoor air with environmental conditions and predicted power draw as inputs. These models were applied in an optimization framework using a genetic algorithm as the optimization solver. The black-box models, developed using extreme gradient boosting (XGB), resulted in a higher accuracy for both the power and temperature change values in the HVAC system and higher accuracy for the power values in the water heater system, due to their ability to infer nonlinear and transient system dynamics that are not incorporated in the simplified linear grey-box models. The accuracies of the grey-box and black-box models are summarized in Table 7.1.

Table 7.1: HVAC and water heater model accuracies

Model	Variable	RMSE	MAE	R^2
Water Heater Grey-Box	Power [W]	149.62	96.14	-
	Temperature [ºC]	0.383	0.170	0.954
Water Heater Black-Box	Power [W]	18.0	13.0	0.973
	Temperature [ºC]	0.897	0.295	0.900
HVAC Grey-Box (EES Data)	Power [W]	2.45	1.90	0.998
	Temperature [ºC]	0.083	0.063	0.980
HVAC Black-Box (EES Data)	Power [W]	1.0	1.0	0.999
	Temperature [ºC]	0.066	0.042	0.988

The genetic algorithm was selected due to its ability to effectively solve the derivative-free, combinatorial optimization problem resulting from the use of the XGB algorithm for black-box modeling. The problem was formulated by adding a penalty function to the objective function, which would generate a penalty value that increased as temperatures deviated from the thermal comfort temperature set point and dead band of each appliance. The use of a penalty function resulted in an unconstrained problem, rather than a constrained problem subject to the constraints defined by the thermal comfort dead band of each appliance. Posing the problem as an unconstrained optimization allowed for the solver to select solutions may be globally optimal but violate the temperature constraints, rendering them infeasible when solved using a constrained method. Constrained methods of optimization may result in infeasible solutions and increased energy consumption, as demonstrated in Section 6.1 and further discussed in Appendix A. Weights were chosen for the penalty function based on the outcome of a parametric study in which weights for the HVAC and water heating system operation were varied and evaluated for the resulting combination of cost savings and thermal comfort.

Results of this study show that energy savings, peak load shifting, and cost savings may be achieved by controlling the water heater and HVAC system of a house, and that

accurate models of these systems may be developed using machine learning methods. Monthly total energy consumption was reduced by up to 12.9% and monthly peak energy consumption reduced by up to 41.3% while maintaining the water reservoir and indoor air temperatures within the designated temperature dead band. However, these percentages do not translate into very large absolute values. The 12.9% total energy reduction seen during the month of August only amounts to a 46.19 kWh reduction and the 41.3% peak energy consumption reduction seen in January translates to a 42.56 kWh reduction. Maximum cost savings for the customer amounted to $6.86 for the month of August and $3.20 for the month of January. The HERS rating of the house studied indicates that it is expected to consume approximately 40% of the energy consumed by a home of typical construction in the same climate, but even adjusting for this results in peak energy reduction of 106.4 kWh, or approximately 3.5 kWh per day. For this amount of peak energy reduction to make a significant contribution in a demand response program, the algorithm must be deployed at a large scale in which the load shifting capability of many houses is aggregated. Application of the algorithm across a region requires the use of generalized system models that may adapt to a range of appliance variations encountered in the field. Developing black-box models of the appliances using machine learning, rather than adapting a physics-based model to each appliance, is desirable because it may allow the optimization algorithm to be applied to water heaters and HVAC systems of any type, from any manufacturer. In summary, three key contributions of this thesis are:

- Accurate models of water heating and HVAC systems may be developed using black-box methods, if the necessary data are provided for training. These

models may be adapted to a variety of appliance types, which would allow for deployment of the algorithm on a large scale,

- Applying a combinatorial optimization solver such as the genetic algorithm allows for the optimization of the derivative-free objective function resulting from the use of black-box models, while also allowing the use of a penalized objective function. This allows for optimization of systems that may be problematic in a constrained format,
- Cost savings for the customer and peak energy savings for the utility are achievable using this optimization algorithm when the objectives are weighted appropriately.

7.1 Recommendations for Future Work

To further investigate the efficacy of the black-box modeling methods used in this study, the XGB algorithm used to model the water heater and HVAC system may be applied to water heaters and HVAC systems for a variety of manufacturers and model numbers. Ideally, this evaluation would be conducted in a laboratory setting in which all environmental factors may be controlled and/or measured. This would allow for a thorough investigation of the application of machine learning methods to systems such as the multi-zone, variable capacity HVAC system installed in the house monitored for this study. The performance data for this system were ultimately replaced in the algorithm by the simulated performance of a single-zone, single-speed system of a similar capacity due to the incompleteness of the field data set. Testing units like this in a controlled setting is necessary to validate the machine learning modeling methods before moving to a field test setting. The modeling methods may also be extended to gas-fired appliances for

optimization scenarios in which they are included. All modeling parameters would be the same, with the exception of power, which would be replaced by natural gas volumetric flow rate, and price, which would need the addition of a natural gas pricing schedule.

The optimization algorithm may be improved by the addition of a feature that allows for the approximation of the Pareto front and the subsequent assignment of objective function weights as part of the solution process. This would be an improvement over the manual, iterative evaluation of Pareto-optimal solutions conducted in this study that resulted in a static set of weights used for all simulations within a given month. The algorithm may also be improved by increasing its computational speed. Optimization with the black-box models required approximately six hours of computing time on a MacBook Pro with a 2.4 GHz Intel i9 processor and 32 GB 2400 MHz DDR4 memory. The algorithm was written in CPython, due to ease of use and the availability of various libraries and toolkits that aided in the algorithm design. Python is an interpreted language and is known to be slower than a compiled language like C++. Converting the script to a different language may result in faster execution. In addition, the genetic algorithm is well-suited to parallel computation, but was executed in serial for this study. Parallelization of the algorithm may result in a shorter computation time.

Additional optimization objectives may also be added to the objective function in the future. For example, the minimization of carbon emissions may be considered. This would involve the evaluation of utility generation mixes in the local area as well as various appliance fuel types. For example, gas appliances may be included in the optimization, with the carbon emissions from gas-fired appliances being integrated into an algorithm that also considers the carbon emissions of local electrical utility power generation when

deciding an optimal control schedule. Houses may also be optimized in aggregate to share generation resources and maximize load shifting potential.

Pricing schemes other than the time-of-use rate, including real-time pricing or critical peak pricing, may also be considered. The use of a pricing schedule may increase participation in utility demand response programs by incentivizing customers with greater cost savings.

APPENDIX A. WATER HEATER HEAT PUMP TEMPERATURE RECOVERY CAPABILITY AND CONSTRAINED OPTIMIZATION

There are instances in which the heat pump of the water heater is incapable of heating the water reservoir to a temperature within the set point dead band during the first 15-minute interval of optimization, which results in the activation of the two electric resistive elements, greatly increasing the electricity consumption and cost. Sample calculations for the water heater temperature recovery functionality using the heat pump, and the heat pump combined with a resistance element, are provided in Table A.1 and Table A.2, respectively. For illustrative purposes, a constant condenser heat duty and a well-mixed water reservoir are assumed. Losses from the water reservoir to the ambient are considered to be negligible. This calculation demonstrates the effect of a high-volume water draw, 0.10 m^3 (20 gallons), during the first timestep of optimization. The set point of the water heater is 49ºC, while the minimum dead band temperature is 43ºC. When

Table A.1: Water heater temperature recovery calculation - heat pump only

Inputs	Equations	Results
$V_{draw} = 0.10\ \text{m}^3$ $\rho_w = 998\ \text{kg m}^{-3}$ $c = 4.186\ \text{kJ kg}^{-1}\ \text{K}^{-1}$ $T_{inlet}(t) =$ 20ºC $T_w(t) =$ 49ºC $\tau = 900\ \text{s}$ $P_{nom,HP} = 0.405\ \text{kW}$ $COP = 3.5$ $U_{HP} = 1$ $V_{tank} = 0.27\ \text{m}^3$	Thermal energy loss due to water draw $Q_{draw} = V_{flow}\rho_w c\left(T_{inlet}(t) - T_w(t)\right)$	–9172.22 kJ
	Thermal energy added by heat pump $Q_{HP} = \tau P_{nom,HP} COP \cdot U_{HP}$	1275.75 kJ
	Temperature at time $t = t + 1$ $T_w(t+1) = T_w(t) + \frac{1}{V_{tank}\rho_w c}\left[Q_{draw} + Q_{HP}\right]$	42.1°C

Table A.2: Water heater temperature recovery calculation - heat pump and electric resistive element

Inputs	Equations	Results
$V_{draw} = 0.10\ \text{m}^3$ $\rho_w = 998\ \text{kg m}^{-3}$ $c = 4.186\ \text{kJ kg}^{-1}\ \text{K}^{-1}$ $T_{inlet}(t) =\ 20^{\circ}\text{C}$ $T_w(t) =\ 49^{\circ}\text{C}$ $\tau = 900\ \text{s}$ $P_{nom,HP} = 0.405\ \text{kW}$ $P_{nom,ER} = 2.5\ \text{kW}$ $COP = 3.5$ $\eta_{ER} = 0.98$ $U_{HP} = 1$ $U_{ER} = 1$	Thermal energy loss due to water draw $Q_{draw} = V_{flow}\rho_w c\left(T_{inlet}(t) - T_w(t)\right)$	−9172.22 kJ
	Thermal energy added by heat pump $Q_{HP} = \tau P_{nom,HP} COP \cdot U_{HP}$	1275.75 kJ
	Thermal energy added by electric resistance elements $Q_{ER} = \tau P_{nom,ER}\eta_{ER}U_{ER}$	2205 kJ
	Temperature at time $t = t + 1$ $T_w(t+1) = T_w(t) + \frac{1}{V_{tank}\rho_w c}\left[Q_{draw} + Q_{HP} + Q_{ER}\right]$	44°C

heating the water using only the heat pump, the water reaches a temperature of 42.1ºC, which is below the minimum set point temperature and would result in an infeasible solution in a constrained problem. Activating one electric resistive element during this time step allows the water heater to reheat to 44ºC.

Example calculations are provided to demonstrate the difference between a constrained and penalized optimization problem using a simple objective function. A contour plot of the problem space is shown in Figure A.1. The global minimum for this function is located at (0,0). The dashed lines represent the objective constraints, with the shaded region representing the region of feasible solutions. The minimum value for the function achievable without violating the constraints is (-2,0), and the objective function result for this location is calculated in Table A.3.

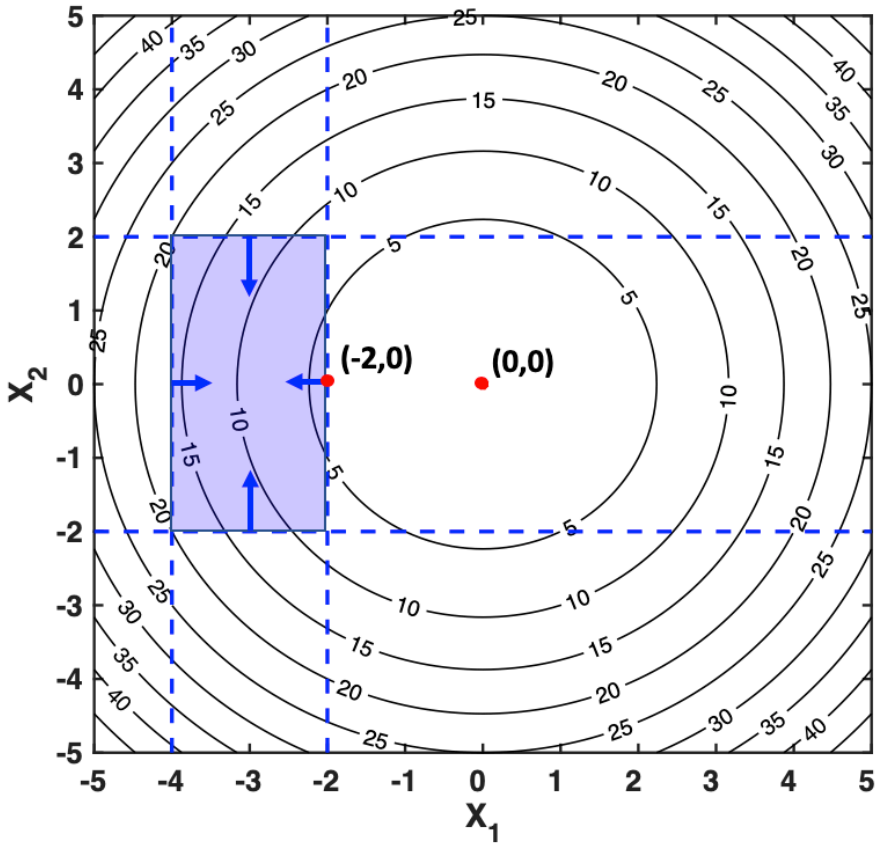

Figure A.1: Optimization region with constraints and global minimum

Converting the constraints to penalties and adding them to the objective function results in an unconstrained problem that may deviate from the region designated as feasible in the constrained case, but suffer a penalty by doing so. The outcome may, however, be an improvement, regardless of the penalty. An example of the use of a penalty function is given in Table A.4.

Table A.3: Constrained optimization objective function

Constrained Optimization		
Inputs	**Equations**	**Results**
Location of constrained minimum: $x_1 = -2$ $x_2 = 0$	Objective: Minimize $f(x_1, x_2) = x_1^2 + x_2^2$	$f(-2,0) = 4$
	Constraints: $-4 \le x_1 \le -2$ $-2 \le x_2 \le 2$	

Table A.4: Penalty function in optimization

Penalty Function in Optimization		
Inputs	**Equations**	**Results**
Location of constrained minimum: $x_1 = 0$ $x_2 = 0$	Objective: Minimize $f(x_1, x_2) = x_1^2 + x_2^2 + p_1 + p_2$	$f(0,0) = 2$
	Penalty functions: $p_1 = \begin{cases} \text{if } x_1 < -4 \lvert x_1 - (-4) \rvert \\ \text{if } x_1 > -2 \; x_1 - (-2) \end{cases}$ $p_2 = \begin{cases} \text{if } x_2 < -2 \lvert x_2 - (-2) \rvert \\ \text{if } x_2 > 2 \quad x_2 - 2 \end{cases}$	

APPENDIX B. UNCERTAINTY PROPAGATION

Uncertainty propagation for the power measurement was calculated by summing the error of each power measurement and prediction for a given interval in quadrature. The same operation is then performed for all 15-minute intervals within the time period. An example calculation of uncertainty propagation is provided in Table B.1. Measurement uncertainties are given in Table B.2.

Table B.1: Power uncertainty calculation

Power Uncertainty		
Inputs	**Equations**	**Results**
$P_{WH,1} = 0.40$ kW $P_{HVAC,1} = 1$ kW $P_{unctrl,1} = 0.5$ kW $P_{PV,1} = 1.5$ kW $P_{WH,2} = 0.40$ kW $P_{HVAC,2} = 1$ kW $P_{unctrl,2} = 0.75$ kW $P_{PV,2} = 1.5$ kW	$P_{grid} = P_{WH} + P_{HVAC} + P_{unctrl} - P_{PV}$	
	$\delta P_{grid,i} = \sqrt{\left(\delta P_{WH,i}\right)^2 + \left(\delta P_{HVAC,i}\right)^2 + \left(\delta P_{unctrl,i}\right)^2 + \left(\delta P_{PV,i}\right)^2}$	$\delta P_{grid,1} = 0.173$ kW $\delta P_{grid,2} = 0.173$ kW
	$\delta P_{grid} = \sqrt{\sum_i^N \left(\delta P_{grid,i}\right)^2}$	$\delta P_{grid} = 0.24$ kW

Table B.2: Measurement uncertainties

Measurement Uncertainties
$\delta P_{WH} = 0.17$ kW
$\delta P_{HVAC} = 0.01$ kW
$\delta P_{unctrl} = 0.02 * P_{unctrl}$
$\delta P_{PV} = 0.02 * P_{PV}$

The uncertainties for the maximum and minimum temperatures reported were the mean absolute error (MAE) of the temperature prediction. The uncertainty of the mean temperature was calculated in the same manner as the power uncertainty.

APPENDIX C. TABULATED RESULTS

August Total Energy (with PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Reduction**
1	7.35	6.89	6
2	12.93	10.83	16
3	10.54	9.10	14
4	15.76	10.75	32
5	15.87	14.5	9
6	10.53	9.14	13
7	13.80	12.09	12
8	12.28	10.91	11
9	12.74	10.58	17
10	13.35	11.11	17
11	11.63	10.21	12
12	13.15	11.28	14
13	14.71	12.85	13
14	12.11	10.98	9
15	14.50	12.67	13
16	10.94	8.76	20
17	12.67	11.19	12
18	13.63	12.45	9
19	15.41	14.22	8
20	13.39	12.43	7
21	9.81	9.06	8
22	11.88	11.72	1
23	9.21	8.32	10
24	8.69	7.62	12
25	13.13	13.01	1
26	9.84	9.59	3
27	9.69	8.83	9
28	9.39	8.29	12
29	11.22	9.44	16
30	9.00	7.30	19
31	-	-	-
Total	**359.15**	**316.12**	**12.0**

August Peak Energy (with PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Difference**
1	0.37	0.24	35
2	3.75	3.38	10
3	0.73	0.29	60
4	1.83	1.99	-9
5	1.63	1.00	39
6	1.99	1.26	37
7	0.26	0.09	65
8	0.49	0.21	57
9	1.15	0.91	21
10	1.86	0.86	54
11	1.95	1.62	17
12	2.71	1.47	46
13	0.61	0.34	44
14	0.15	0.00	100
15	1.24	0.65	48
16	0.40	0.16	60
17	0.79	0.44	44
18	3.21	2.47	23
19	3.81	2.78	27
20	1.86	1.09	41
21	0.43	0.25	42
22	0.48	0.37	23
23	0.15	0.08	47
24	1.26	0.48	62
25	2.94	2.37	19
26	0.93	0.74	20
27	0.61	0.29	52
28	0.55	0.18	67
29	0.83	0.46	45
30	0.52	0.22	58
31	-	-	-
Total	**39.49**	**26.69**	**32**

August Total Cost (with PV)			
Day	**Baseline [$]**	**Optimized [$]**	**% Difference**
1	0.62	0.57	8
2	1.56	1.34	14
3	0.92	0.75	18
4	1.49	1.13	24
5	1.47	1.27	14
6	1.11	0.89	20
7	1.10	0.95	14
8	1.02	0.87	15
9	1.16	0.95	18
10	1.31	0.99	24
11	1.19	1.03	13
12	1.42	1.09	23
13	1.23	1.04	15
14	0.96	0.85	11
15	1.31	1.08	18
16	0.91	0.70	23
17	1.10	0.93	15
18	1.53	1.33	13
19	1.76	1.52	14
20	1.31	1.12	15
21	0.82	0.74	10
22	0.99	0.96	3
23	0.73	0.65	11
24	0.86	0.66	23
25	1.46	1.36	7
26	0.90	0.85	6
27	0.84	0.72	14
28	0.81	0.67	17
29	0.99	0.80	19
30	0.77	0.60	22
31	-	-	-
Total	**33.66**	**28.43**	**16**

January Total Energy (with PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Reduction**
1	22.10	23.98	-8
2	26.34	27.37	-4
3	32.65	33.37	-2
4	19.53	21.68	-11
5	30.08	28.85	4
6	18.14	17.62	3
7	15.46	15.15	2
8	16.46	17.10	-4
9	18.24	19.48	-7
10	19.59	18.41	6
11	26.59	23.07	13
12	29.82	29.05	3
13	28.75	27.31	5
14	24.14	23.42	3
15	23.25	21.62	7
16	18.31	16.92	8
17	23.86	21.97	8
18	17.64	17.01	4
19	21.12	22.88	-8
20	34.41	30.41	12
21	22.17	19.71	11
22	19.50	16.46	16
23	23.13	21.00	9
24	-	-	-
25	-	-	-
26	-	-	-
27	25.47	24.69	3
28	18.43	16.98	8
29	24.18	22.25	8
30	21.70	19.84	9
31	-	-	-
Total	**621.05**	**597.57**	**4**

January Peak Energy (with PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Reduction**
1	2.80	1.23	6
2	3.96	2.71	16
3	4.31	3.01	14
4	3.88	1.81	32
5	2.86	2.03	9
6	2.73	2.09	13
7	2.64	1.87	12
8	3.64	1.57	11
9	2.05	1.86	17
10	3.94	2.03	17
11	3.79	1.95	12
12	5.03	3.76	14
13	3.46	2.23	13
14	4.54	2.88	9
15	4.89	3.11	13
16	3.94	2.05	20
17	4.17	2.50	12
18	4.01	2.61	9
19	2.81	1.53	8
20	4.42	2.60	7
21	5.01	2.43	8
22	3.16	2.14	1
23	4.50	1.92	10
24	-	-	-
25	-	-	-
26	-	-	-
27	3.28	2.69	9
28	3.28	1.73	12
29	4.47	2.70	16
30	4.68	2.47	19
31	-	-	-
Total	**102.26**	**61.51**	**40**

January Total Cost (with PV)			
Day	**Baseline [$]**	**Optimized [$]**	**% Difference**
1	7.05	7.50	-6
2	2.11	2.17	-3
3	2.61	2.64	-1
4	1.59	1.71	-8
5	2.38	2.27	5
6	1.45	1.40	4
7	1.25	1.21	3
8	1.34	1.35	-1
9	1.45	1.54	-6
10	1.59	1.46	8
11	2.13	1.82	14
12	2.40	2.32	4
13	2.29	2.15	6
14	1.95	1.87	5
15	1.89	1.73	9
16	1.49	1.35	10
17	1.93	1.75	9
18	1.44	1.37	5
19	1.69	1.80	-7
20	2.75	2.40	13
21	1.81	1.57	13
22	1.57	1.31	16
23	1.88	1.66	12
24	-	-	-
25	-	-	-
26	-	-	-
27	2.03	1.96	4
28	1.49	1.35	10
29	1.96	1.77	9
30	1.77	1.58	11
31	-	-	-
Total	**55.28**	**53.00**	**4**

August Total Energy (no PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Reduction**
1	13.92	13.35	4%
2	23.49	22.30	5%
3	21.20	20.24	5%
4	24.83	22.97	7%
5	35.57	33.65	5%
6	22.50	22.03	2%
7	25.63	23.95	7%
8	25.04	23.09	8%
9	25.73	24.57	5%
10	-	-	-
11	23.82	22.32	6%
12	28.67	27.12	5%
13	26.38	23.99	9%
14	22.93	21.67	5%
15	25.99	23.41	10%
16	20.17	17.55	13%
17	22.89	21.40	7%
18	26.14	25.81	1%
19	29.34	28.95	1%
20	25.23	24.90	1%
21	17.47	17.25	1%
22	20.11	20.11	0%
23	17.36	17.02	2%
24	19.80	18.62	6%
25	26.45	26.13	1%
26	30.12	30.22	0%
27	20.24	19.28	5%
28	17.99	17.25	4%
29	20.60	18.94	8%
30	22.14	20.95	5%
31	-	-	-
Total	**681.75**	**649.04**	**5%**

August Peak Energy (no PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Difference**
1	3.51	2.98	15
2	9.71	8.92	8
3	6.66	5.27	21
4	10.29	8.92	13
5	12.15	10.91	10
6	8.35	7.29	13
7	7.03	5.59	20
8	6.09	5.34	12
9	5.46	5.13	6
10	-	-	-
11	7.83	6.68	15
12	12.81	11.41	11
13	6.88	5.74	17
14	4.83	4.04	16
15	7.45	5.79	22
16	4.93	3.78	23
17	5.92	4.79	19
18	9.41	8.44	10
19	11.57	10.52	9
20	9.64	8.82	9
21	4.30	3.77	12
22	4.36	3.65	16
23	3.54	3.21	9
24	7.93	6.04	24
25	9.45	8.80	7
26	12.13	12.16	0
27	5.77	5.00	13
28	4.44	3.76	15
29	5.04	3.79	25
30	5.28	4.37	17
31	-	-	-
Total	**212.76**	**184.91**	**13**

August Total Cost (no PV)			
Day	**Baseline [$]**	**Optimized [$]**	**% Difference**
1	1.60	1.48	8
2	3.27	3.06	6
3	2.64	2.35	11
4	3.46	3.11	10
5	4.57	4.23	7
6	2.99	2.79	7
7	3.03	2.69	11
8	2.85	2.58	9
9	2.81	2.67	5
10	-	-	-
11	3.01	2.72	10
12	4.13	3.81	8
13	3.07	2.71	12
14	2.49	2.28	8
15	3.12	2.68	14
16	2.30	1.92	17
17	2.65	2.37	11
18	3.43	3.26	5
19	4.00	3.81	5
20	3.39	3.25	4
21	1.99	1.90	5
22	2.21	2.10	5
23	1.87	1.80	4
24	2.72	2.34	14
25	3.46	3.34	3
26	4.14	4.16	0
27	2.43	2.24	8
28	2.06	1.90	8
29	2.35	2.03	14
30	2.50	2.27	9
31	-	-	-
Total	**33.66**	**28.43**	**16**

January Total Energy (no PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Reduction**
1	26.45	25.97	2
2	29.76	31.04	-4
3	36.95	37.67	-2
4	24.10	26.27	-9
5	41.09	41.39	-1
6	28.32	28.59	-1
7	22.27	21.05	5
8	22.63	22.92	-1
9	26.29	28.55	-9
10	30.51	30.28	1
11	37.31	34.87	7
12	35.47	34.44	3
13	32.67	33.09	-1
14	27.67	26.91	3
15	28.92	26.62	8
16	26.51	26.21	1
17	28.47	25.39	11
18	22.66	22.25	2
19	24.29	24.47	-1
20	45.57	42.18	7
21	38.45	35.74	7
22	27.86	24.09	14
23	25.90	23.33	10
24	-	-	-
25	-	-	-
26	-	-	-
27	34.81	33.80	3
28	27.24	25.43	7
29	35.48	33.50	6
30	33.57	31.10	7
31	-	-	-
Total	**821.25**	**797.16**	**3**

January Peak Energy (no PV)			
Day	**Baseline [kWh]**	**Optimized [kWh]**	**% Difference**
1	3.08	1.48	52
2	4.24	2.96	30
3	4.51	3.29	27
4	4.34	2.26	48
5	4.00	3.06	23
6	3.80	3.14	17
7	3.65	2.84	22
8	4.27	3.01	29
9	2.86	2.77	3
10	5.19	2.98	43
11	11.94	10.55	12
12	5.37	4.11	23
13	3.85	3.23	16
14	4.77	3.10	35
15	5.83	3.09	47
16	5.33	3.80	29
17	4.44	2.76	38
18	4.45	3.03	32
19	2.98	1.70	43
20	18.97	18.33	3
21	7.92	4.59	42
22	4.46	2.75	38
23	4.63	2.05	56
24	-	-	-
25	-	-	-
26	-	-	-
27	4.65	3.02	35
28	5.42	3.29	39
29	5.05	3.28	35
30	6.83	4.23	38
31	-	-	-
Total	**146.82**	**104.71**	**29**

January Total Cost (no PV)			
Day	**Baseline [$]**	**Optimized [$]**	**% Difference**
1	2.10	2.03	3
2	2.38	2.46	-3
3	2.94	2.97	-1
4	1.95	2.07	-6
5	3.25	3.26	0
6	2.26	2.27	0
7	1.79	1.68	6
8	1.83	1.83	0
9	2.09	2.26	-8
10	2.98	2.75	8
11	2.98	2.75	8
12	2.85	2.74	4
13	2.60	2.62	-1
14	2.23	2.14	4
15	2.35	2.12	10
16	2.15	2.10	2
17	2.29	2.02	12
18	1.84	1.78	3
19	1.94	1.92	1
20	3.61	3.32	8
21	3.13	2.85	9
22	2.21	2.10	5
23	2.09	1.84	12
24	-	-	-
25	-	-	-
26	-	-	-
27	2.78	2.67	4
28	2.21	2.03	8
29	2.84	2.65	7
30	2.73	2.49	9
31	-	-	-
Total	**66.42**	**63.73**	**4**

REFERENCES

Adika, C. O. and L. Wang (2014). "Demand-Side Bidding Strategy for Residential Energy Management in a Smart Grid Environment," *IEEE Transactions on Smart Grid*. Vol. 5(4), pp. 1724-1733 DOI: 10.1109/TSG.2014.2303096.

Ahmad, S., M. Naeem and A. Ahmad (2019). "Unified Optimization Model for Energy Management in Sustainable Smart Power Systems," *International Transactions on Electrical Energy Systems*. Vol. n/a(n/a), p. e12144 DOI: 10.1002/2050-7038.12144.

Ahmad, T., H. Chen, Y. Guo and J. Wang (2018). "A Comprehensive Overview on the Data Driven and Large Scale Based Approaches for Forecasting of Building Energy Demand: A Review," *Energy and Buildings*. Vol. 165, pp. 301-320 DOI: https://doi.org/10.1016/j.enbuild.2018.01.017.

Alajmi, A. and J. Wright (2014). "Selecting the Most Efficient Genetic Algorithm Sets in Solving Unconstrained Building Optimization Problem," *International Journal of Sustainable Built Environment*. Vol. 3(1), pp. 18-26 DOI: https://doi.org/10.1016/j.ijsbe.2014.07.003.

Amarasinghe, K., D. L. Marino and M. Manic (2017). "Deep Neural Networks for Energy Load Forecasting," *Proceedings of 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*.

Anvari-Moghaddam, A., H. Monsef and A. Rahimi-Kian (2015). "Optimal Smart Home Energy Management Considering Energy Saving and a Comfortable Lifestyle," *IEEE Transactions on Smart Grid*. Vol. 6(1), pp. 324-332 DOI: 10.1109/TSG.2014.2349352.

Asadinejad, A., A. Rahimpour, K. Tomsovic, H. Qi and C.-f. Chen (2018). "Evaluation of Residential Customer Elasticity for Incentive Based Demand Response Programs," *Electric Power Systems Research*. Vol. 158, pp. 26-36 DOI: https://doi.org/10.1016/j.epsr.2017.12.017.

Bakker, V., M. G. C. Bosman, A. Molderink, J. Hurink and G. J. M. Smit (2010). *Demand Side Load Management Using a Three Step Optimization Methodology*.

Bamdad, K., M. E. Cholette, L. Guan and J. Bell (2017). "Ant Colony Algorithm for Building Energy Optimisation Problems and Comparison with Benchmark Algorithms," *Energy and Buildings*. Vol. 154, pp. 404-414 DOI: https://doi.org/10.1016/j.enbuild.2017.08.071.

Bayram, I. S. and T. S. Ustun (2017). "A Survey on Behind the Meter Energy Management Systems in Smart Grid," *Renewable and Sustainable Energy Reviews*. Vol. 72, pp. 1208-1232 DOI: https://doi.org/10.1016/j.rser.2016.10.034.

Beaudin, M. and H. Zareipour (2015). "Home Energy Management Systems: A Review of Modelling and Complexity," *Renewable and Sustainable Energy Reviews*. Vol. 45, pp. 318-335 DOI: https://doi.org/10.1016/j.rser.2015.01.046.

Blum, C. and A. Roli (2003). "Mateheuristics in Combinatorial Optimization: Overview and Conceptual Comparison," *ACM Computing Surveys*. Vol. 35, pp. 268-308.

Borenstein, S., M. Jaske and A. Ros (2002). "Dynamic Pricing, Advanced Metering, and Demand Response in Electricity Markets," University of California Energy Institute, Berkeley, CA.

Chavali, P., P. Yang and A. Nehorai (2014). "A Distributed Algorithm of Appliance Scheduling for Home Energy Management System," *IEEE Transactions on Smart Grid*. Vol. 5(1), pp. 282-290 DOI: 10.1109/TSG.2013.2291003.

Chen, H., Y. Li, R. H. Y. Louie and B. Vucetic (2014). "Autonomous Demand Side Management Based on Energy Consumption Scheduling and Instantaneous Load Billing: An Aggregative Game Approach," *IEEE Transactions on Smart Grid*. Vol. 5(4), pp. 1744-1754 DOI: 10.1109/TSG.2014.2311122.

Chen, T. G., Carlos (2016). "Xgboost: A Scalable Tree Boosting System," *Proceedings of 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA.

Cui, B., C. Fan, J. Munk, N. Mao, F. Xiao, J. Dong and T. Kuruganti (2019). "A Hybrid Building Thermal Modeling Approach for Predicting Temperatures in Typical, Detached, Two-Story Houses," *Applied Energy*. Vol. 236, pp. 101-116 DOI: https://doi.org/10.1016/j.apenergy.2018.11.077.

CVX Research Inc. (2012). *Cvx: Matlab Software for Disciplined Convex Programming*, Version 2.0.

Deilami, S., A. S. Masoum, P. S. Moses and M. A. S. Masoum (2011). "Real-Time Coordination of Plug-in Electric Vehicle Charging in Smart Grids to Minimize Power Losses and Improve Voltage Profile," *IEEE Transactions on Smart Grid*. Vol. 2(3), pp. 456-467 DOI: 10.1109/TSG.2011.2159816.

Denholm, P., R. Margolis and J. Milford (2008). "Production Cost Modeling for High Levels of Photovoltaics Penetration," National Renewable Energy Lab. (NREL), Golden, CO, USA.

Digalakis, J. and K. G. Margaritis (2001). "On Benchmarking Functions for Genetic Algorithm," *International Journal of Computer Mathematics*. Vol. 77 DOI: 10.1080/00207160108805080.

DOE (2006). "Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them: A Report to the United States Congress Pursuant to Section

1252 of the Energy Policy Act of 2005," U.S. Department of Energy, Washington D.C.

EERE (2019). Climate Zones, Access Date: DOI.

EIA (2019a). "Annual Electric Power Industry Report," United States Energy Information Association, Washington, D.C.

EIA (2019b). "Annual Energy Outlook 2019," U.S. Energy Information Administration, Washington D.C.

Faruqui, A. S., Sanem (2009). "Household Response to Dynamic Pricing of Electricity: A Survey of the Experimental Evidence," The Brattle Group.

FERC (2018). "Assessment of Demand Response and Advanced Metering," Federal Energy Regulatory Commission, Washington, D.C.

Golmohamadi, H., R. Keypour, B. Bak-Jensen and J. R. Pillai (2019a). "A Multi-Agent Based Optimization of Residential and Industrial Demand Response Aggregators," *International Journal of Electrical Power & Energy Systems*. Vol. 107, pp. 472-485 DOI: https://doi.org/10.1016/j.ijepes.2018.12.020.

Golmohamadi, H., R. Keypour, B. Bak-Jensen and J. Radhakrishna Pillai (2019b). "Optimization of Household Energy Consumption Towards Day-Ahead Retail Electricity Price in Home Energy Management Systems," *Sustainable Cities and Society*. Vol. 47, p. 101468 DOI: https://doi.org/10.1016/j.scs.2019.101468.

Grant, M. and S. Boyd (2013). *Cvx: Matlab Software for Disciplined Convex Programming*, Version 2.0 beta.

Hendron, R. and C. Engebrecht (2010). "Building America House Simulation Protocols," United States Department of Energy, Oak Ridge, TN.

ISO-NE (2020). *Iso New England Real-Time Maps and Charts*. ISO New England Inc.

Jebari, K. (2013). "Selection Methods for Genetic Algorithms," *International Journal of Emerging Sciences*. Vol. 3, pp. 333-344.

Jin, X., K. Baker, D. Christensen and S. Isley (2017). "Foresee: A User-Centric Home Energy Management System for Energy Efficiency and Demand Response," *Applied Energy*. Vol. 205, pp. 1583-1595 DOI: https://doi.org/10.1016/j.apenergy.2017.08.166.

Klein, S. A. (2018). *Engineering Equation Solver (EES)*, Version 10.644. F-Chart Software, Madison, USA.

Konak, A., D. W. Coit and A. E. Smith (2006). "Multi-Objective Optimization Using Genetic Algorithms: A Tutorial," *Reliability Engineering & System Safety*. Vol. 91(9), pp. 992-1007 DOI: https://doi.org/10.1016/j.ress.2005.11.018.

Lampropoulos, I., T. Alskaif, J. Blom and W. van Sark (2019). "A Framework for the Provision of Flexibility Services at the Transmission and Distribution Levels through Aggregator Companies," *Sustainable Energy, Grids and Networks*. Vol. 17, p. 100187 DOI: https://doi.org/10.1016/j.segan.2018.100187.

Laugs, G. A. H., R. M. J. Benders and H. C. Moll (2020). "Balancing Responsibilities: Effects of Growth of Variable Renewable Energy, Storage, and Undue Grid Interaction," *Energy Policy*. Vol. 139, p. 111203 DOI: https://doi.org/10.1016/j.enpol.2019.111203.

Leitão, J., P. Gil, B. Ribeiro and A. Cardoso (2018). "Improving Household's Efficiency Via Scheduling of Water and Energy Appliances," *Proceedings of 2018 13th APCA International Conference on Automatic Control and Soft Computing (CONTROLO)*.

Leitão, J., P. Gil, B. Ribeiro and A. Cardoso (2020). "A Survey on Home Energy Management," *IEEE Access*. Vol. 8, pp. 5699-5722.

Li, N., L. Chen and S. H. Low (2011). "Optimal Demand Response Based on Utility Maximization in Power Networks," *Proceedings of 2011 IEEE Power and Energy Society General Meeting*.

Liao, C.-L., S.-J. Lee, Y.-S. Chiou, C.-R. Lee and C.-H. Lee (2018). "Power Consumption Minimization by Distributive Particle Swarm Optimization for Luminance Control and Its Parallel Implementations," *Expert Systems with Applications*. Vol. 96, pp. 479-491 DOI: https://doi.org/10.1016/j.eswa.2017.11.002.

Liu, Y., C. Yuen, S. Huang, N. U. Hassan, X. Wang and S. Xie (2014). "Peak-to-Average Ratio Constrained Demand-Side Management with Consumer's Preference in Residential Smart Grid," *IEEE Journal of Selected Topics in Signal Processing*. Vol. 8(6), pp. 1084-1097 DOI: 10.1109/JSTSP.2014.2332301.

Lokeshgupta, B. and S. Sivasubramani (2019). "Multi-Objective Home Energy Management with Battery Energy Storage Systems," *Sustainable Cities and Society*. Vol. 47, p. 101458 DOI: https://doi.org/10.1016/j.scs.2019.101458.

Marzband, M., H. Alavi, S. S. Ghazimirsaeid, H. Uppal and T. Fernando (2017). "Optimal Energy Management System Based on Stochastic Approach for a Home Microgrid with Integrated Responsive Load Demand and Energy Storage," *Sustainable Cities and Society*. Vol. 28, pp. 256-264 DOI: https://doi.org/10.1016/j.scs.2016.09.017.

Mathieu, J. L. (2012). *Modeling, Analysis, and Control of Demand Response Resources*. Doctor of Philosophy Thesis, Department of Mechanical Engineering, University of California, Berkeley, Berkeley, CA.

MathWorks (2020). *MATLAB Regression Learner App*, Version R2020a. The MathWorks, Inc., USA.

Miceli, R. (2013). "Energy Management and Smart Grids," *Energies*. Vol. 6, pp. 2262-2290.

Nguyen, D. and L. Le (2014). "Joint Optimization of Electric Vehicle and Home Energy Scheduling Considering User Comfort Preference," *Smart Grid, IEEE Transactions on*. Vol. 5, pp. 188-199 DOI: 10.1109/TSG.2013.2274521.

Palensky, P. and D. Dietrich (2011). "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads," *IEEE Transactions on Industrial Informatics*. Vol. 7(3), pp. 381-388 DOI: 10.1109/TII.2011.2158841.

Rasheed, M., N. Javaid, A. Ahmad, Z. Khan, U. Qasim and N. Alrajeh (2015). "An Efficient Power Scheduling Scheme for Residential Load Management in Smart Homes," *Applied Sciences*. Vol. 5 DOI: 10.3390/app5041134.

RESNET (2019). Understanding the HERS Index, Access Date: October 2018, DOI: https://www.hersindex.com/hers-index/understanding-hers-index/.

Saber, A. Y. and G. K. Venayagamoorthy (2011). "Plug-in Vehicles and Renewable Energy Sources for Cost and Emission Reductions," *IEEE Transactions on Industrial Electronics*. Vol. 58(4), pp. 1229-1238 DOI: 10.1109/TIE.2010.2047828.

Samadi, P., H. Mohsenian-Rad, V. W. S. Wong and R. Schober (2013). "Tackling the Load Uncertainty Challenges for Energy Consumption Scheduling in Smart Grid," *IEEE Transactions on Smart Grid*. Vol. 4(2), pp. 1007-1016 DOI: 10.1109/TSG.2012.2234769.

Shareef, H., M. S. Ahmed, A. Mohamed and E. Al Hassan (2018). "Review on Home Energy Management System Considering Demand Responses, Smart Technologies, and Intelligent Controllers," *IEEE Access*. Vol. 6.

Sortomme, E., M. M. Hindi, S. D. J. MacPherson and S. S. Venkata (2011). "Coordinated Charging of Plug-in Hybrid Electric Vehicles to Minimize Distribution System Losses," *IEEE Transactions on Smart Grid*. Vol. 2(1), pp. 198-205 DOI: 10.1109/TSG.2010.2090913.

Sparn, B., L. Earle, D. Christensen, J. Maguire, E. Wilson and C. E. Hancock (2013). *Field Monitoring Protocol: Heat Pump Water Heaters*. DOE. National Renewable Energy Laboratory, Golden, CO.

Sparn, B., K. Hudon and D. Christensen (2014). *Laboratory Performance Evaluation of Residential Integrated Heat Pump Water Heaters*. DOE. National Renewable Energy Laboratory, Golden, CO.

Verhelst, C., F. Logist, J. Van Impe and L. Helsen (2012). "Study of the Optimal Control Problem Formulation for Modulating Air-to-Water Heat Pumps Connected to a Residential Floor Heating System," *Energy and Buildings*. Vol. 45, pp. 43-53 DOI: https://doi.org/10.1016/j.enbuild.2011.10.015.

Wu, X., X. Hu, S. Moura, X. Yin and V. Pickert (2016). "Stochastic Control of Smart Home Energy Management with Plug-in Electric Vehicle Battery Energy Storage and Photovoltaic Array," *Journal of Power Sources*. Vol. 333, pp. 203-212 DOI: https://doi.org/10.1016/j.jpowsour.2016.09.157.

Xu, Z., R. Diao, S. Lu, J. Lian and Y. Zhang (2014). "Modeling of Electric Water Heaters for Demand Response: A Baseline Pde Model," *IEEE Transactions on Smart Grid*. Vol. 5(5), pp. 2203-2210 DOI: 10.1109/TSG.2014.2317149.

Yeniay, Ö. (2005). "Penalty Function Methods for Constrained Optimization with Genetic Algorithms," *Mathematical and Computational Applications*. Vol. 10, pp. 45-56.